AD-A269 025

OPTICAL COMPUTING

Sponsored by
Optical Society of America

1993 TECHNICAL DIGEST
SERIES VOLUME 7

MARCH 16–19, 1993
PALM SPRINGS,
CALIFORNIA

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

2

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
		Final Report 15 Feb 93 - 14 Jan 94

4. TITLE AND SUBTITLE

Optical Computing

Organization of the 1993 Photonics Science topical Meetings

5. FUNDING NUMBERS

F49620-93-1-0181

6. AUTHOR(S)

Dr Jarus W Quinn

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Optical Society of America
2010 Massachusetts Avenue NW
Washington DC 20036

8. PERFORMING ORGANIZATION REPORT NUMBER

AFOSR-TR- 93 0643

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFOSR/NE
110 Duncan Avenue Suite B115
Bolling AFB DC 20332-0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

2301/DS

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

UNLIMITED

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The following symposium was held

Advanced Solid State Lasers
Compact Blue-Green Lasers
Integrated Photonics Research
Nonlinear Guide-Wave Optics
Optical Amplifiers & Their Applications
Optical Design for Photonics
Photonics in Switching
Quantum Optoelectronics
Shortwavelength: Physics With Intense Laser Pulses
Soft X-Ray Protection Lithography
Ultrafast Electronics & Optoelectronics
Optical Computing
Spatial Light Modulators

93-20349

14. SUBJECT TERMS

93 8 31 069

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASS	UNCLASS	UNCLASS	UL

Optical Computing

Summaries of papers presented at the Optical Computing Topical Meeting

March 16-19, 1993
Palm Springs, California

1993 Technical Digest Series
Volume 7

CONFERENCE EDITION

Sponsored by
Optical Society of America

Optical Society of America
2010 Massachusetts Avenue, NW
Washington, DC 20036-1023

Articles in this publication may be cited in other publications. In order to facilitate access to the original publication source, the following form for the citation is suggested:

Name of Author(s), "Title of Paper," in Optical Computing Technical Digest, 1993 (Optical Society of America, Washington, D.C., 1993), Vol. 7, pp. xx-xx.

ISBN Number

Conference Edition	1-55752-289-8
Postconference Edition	1-55752-290-1
(Note: Postconference Edition includes postdeadline papers.)	
1993 Technical Digest Series	1-55752-317-7

Library of Congress Catalog Card Number

Conference Edition	92-62844
Postconference Edition	92-62843

Printed in U.S.A.

CONTENTS

OPTICAL COMPUTING
TECHNICAL PROGRAM COMMITTEE

B. Keith Jenkins, *General Chair, University of Southern California*

H. Scott Hinton, *Program Chair, AT&T Bell Laboratories*

P. J. Anthony, *AT&T Bell Laboratories*

Mark Bendett, *IMRA America, Inc.*

K. H. Brenner, *University of Erlangen-Nuernberg, Federal Republic of Germany*

Pierre Chavel, *Institut d'Optique, France*

Jacek Chrostowski, *National Research Council, Canada*

Larry A. Coldren, *University of California–Santa Barbara*

Matthew W. Derstine, *Boeing Electronics Co.*

S. C. Esener, *University of California–San Diego*

I. Glaser, *Tel Aviv University, Israel*

J. W. Goodman, *Stanford University*

J. P. Huignard, *Thomson-CSF, France*

Yoshiki Ichioka, *Osaka University, Japan*

S. Ishihara, *Electrotechnical Laboratory, Japan*

Fedor V. Karpushko, *BSSR Academy of Sciences, Byelorussia*

K. Kasahara, *NEC Corporation, Japan*

Soo Young Lee, *Korea Advanced Institute, Korea*

Richard A. Linke, *NEC Research Institute*

Richard J. Michalak, *Griffiss Air Force Base*

J. E. Midwinter, *University College London, U.K.*

Paul R. Prucnal, *Princeton University*

Demetri Psaltis, *California Institute of Technology*

Bernard H. Soffer, *Hughes Research Laboratories*

F. A. P. Tooley, *Heriot-Watt University, U.K.*

Kelvin H. Wagner, *University of Colorado*

Cardinal Warde, *Massachusetts Institute of Technology*

GRAND BALLROOM EAST

2:00 pm–5:30 pm
JTuC, JOINT OPTICAL COMPUTING/PHOTONICS IN SWITCHING/SPATIAL LIGHT MODULATORS PLENARY SESSION
B. Keith Jenkins, *University of Southern California*
Joseph W. Goodman, *Stanford University, Presiders*

2:00 pm (Plenary)
JTuC1 Extended generalized shuffle networks, G. W. Richards, *AT&T Bell Laboratories.* This talk discusses how extended generalized shuffle networks are useful in dealing with various constraints that are encountered when implementing switching networks with photonic technology. **(p. 2)**

2:45 pm (Plenary)
JTuC2 ATM objectives and requirements for next-generation networks, Kai Y. Eng, *AT&T Bell Laboratories.* An overview of ATM networking is described with emphasis on virtual path transport architectures and features. SDH termination, ATM cell processing, and routing are discussed. **(p. 6)**

3:30 pm–4:00 pm COFFEE BREAK

4:00 pm (Plenary)
JTuC3 Photonics in switching: European systems demonstrators and the long-term perspective, Lars Thylèn, *Royal Institute of Technology, Sweden.* Status in the area of photonics-in-switching in Europe, highlighted by systems demonstrators, is reviewed, and the long-term perspective for photonics-in-switching is discussed. **(p. 10)**

4:45 pm (Plenary)
JTUC4 Transition from optical interconnections to optical computing, Richard C. Williamson, *MIT Lincoln Laboratory.* Optical interconnections will be a foot in the door to the next generation of computer systems and provide an evolutionary path toward the use of optics at finer scales. **(p. 15)**

GRAND BALLROOM CENTER

6:30 pm–8:00 pm
OTuA, OPTICAL COMPUTING POSTERS: 1/ CONFERENCE RECEPTION

OTuA1 Concept and implementation of the fractional Fourier transform, Adolf W. Lohmann, *Physikalisches Institut der Univ., Germany.* When a signal $\mu(x)$ is Fourier transformed, its Wigner distribution is rotated by 90°. Less than 90° means fractional transform. Two transform experiments are proposed. **(p. 18)**

OTuA2 Iterative design of computer-generated Fresnel holograms for free-space optical interconnections, Bernard C. Kress, Sing H. Lee, *UC-San Diego.* An iterative algorithm to design CGHs for optical interconnections is presented and compared to noniterative methods. It is applied to twin-butterfly interconnection architecture for multiprocessor parallel computing. **(p. 22)**

OTuA3 Crosstalkless incoherent optical/electronic hybrid associative memory system, Masaki Taniguchi, Yoshiki Ichioka, *Osaka Univ., Japan;* Katsunori Matsuoka, *Government Industrial Research Institute, Osaka, Japan.* By using multiple-object discriminant filters calculated by simulated annealing algorithm, crosstalkless association results are obtained in incoherent optical/electronic hybrid associative memory system. **(p. 26)**

OTuA4 Reflective block optics for optical computing systems, Daisuke Miyazaki, Jun Tanida, Yoshiki Ichioka, *Osaka Univ., Japan.* A new concept of the reflective block optics for optical digital computing systems is proposed. A discrete correlator is designed, and basic experiments are executed to verify the concept. **(p. 30)**

OTuA5 Cellular two-layer logic array and optoelectronic implementation, Liren Liu, Zibei Zhang, Xuejun Zhang, Jun Zheng, *Shanghai Institute of Optics and Fine Mechanics, China.* A new architecture of cellular-morphologic two-layer logic array is proposed for binary and grey-level image processing and for image transformation. An optoelectronic system is developed. **(p. 34)**

OTuA6 Experimental investigation of the tolerance of S-SEEDs and L-SEEDs, Frank A. P. Tooley, *Heriot-Watt Univ., UK;* Anthony L. Lentine, Sue Wakelin, *AT&T Bell Laboratories;* Alex Wachlowski, Dominic Goodwill, Doug Baillie, Kevin Simpson, *Alcatel, Austria.* Experimental results will be presented of measurements of the tolerance to spatial noise of symmetric-SEEDs and logic-SEEDs. These are compared with predictions from a model. **(p. 38)**

OTuA7 Paper withdrawn.

OTuA8 Neural network router for optical interconnection networks, C. Lee Giles, Mark W. Goudreau, *NEC Research Institute.* We propose a neural network routing methodology that can generate control bits for an optical multistage interconnection network and suggest an optical implementation. **(p. 42)**

OTuA9 Design of a microlens-based total interconnection for optical neural networks, P. C. H. Poon, D. R. Selviah, J. E. Midwinter, *Univ. College London, U.K.;* D. Daly, *National Physical Laboratory, U.K.;* M. G. Robinson, *Sharp Labs. Europe, U.K.* Two system designs are analyzed, and two types of microlens are characterized to design a microlens based total interconnection for optical neural networks. **(p. 46)**

OTuA10 High-speed database processing on an optical content addressable parallel processor (OCAPP), Ahmed Louri, James A. Hatch, Jr., *Univ. Arizona.* This paper describes the use of the OCAPP for efficient and high-speed database processing. **(p. 50)**

OTuA11 Volume holographic storage and processing for large databases, Jeff Brown, L. J. Irakliotis, Pericles A. Mitkas, *Colorado State Univ.* A volume holographic database system, that operates in a record-parallel mode and allows associative and address-based data access and retrieval, is described. **(p. 54)**

WEDNESDAY, MARCH 17, 1993

MESQUITE ROOM

8:15 am–8:30 am
OPENING REMARKS
B. K. Jenkins, *University of Southern California*

MESQUITE ROOM

8:30 am–10:00 am
OWA, WAVELENGTH DOMAIN PROCESSING
Demetri Psaltis, *California Institute of Technology, Presider*

8:30 am (Invited)
OWA1 Spectral hole burning, holographic storage and processing, Urs P. Wild, Stefan Bernet, Stefan Altner, Eric S. Maniloff, Alois Renn, *Swiss Federal Institute of Technology, Switzerland.* Spectral hole burning allows for storage and simultaneous processing of data. Using frequency and electric field multiplexing this capability and its applications will be discussed. **(p. 60)**

9:00 am
OWA2 Multiwavelength logic gates with high computational efficiency, D. J. Blumenthal, J. R. Sauer, *Univ. Colorado, Boulder.* Two classes of logic gate which process information based on the presence and absence of optical power at discrete wavelengths are introduced. These gates exhibit computational efficiencies many orders of magnitude greater than Boolean logic. An example processor is investigated and advantages and limitations are discussed. **(p. 64)**

9:20 am
OWA3 Multiwavelength optical half adder, Pochi Yeh, Scott Campbell, Shaomin Zhou, *UC–Santa Barbara.* We propose and demonstrate spectrally parallel CARRY and SUM operations via optical four-wave mixing in photorefractive media to achieve a multiwavelength optical half adder. **(p. 68)**

9:40 am
OWA4 Wavelength multiplexed computer-generated volume holography, Joseph Rosen, *USAF Rome Laboratories;* Mordechai Segev, Amnon Yariv, *California Institute of Technology.* We demonstrate recording and reconstruction of multiple computer-generated, wavelength multiplexed, volume holograms, in a holographic storage medium. The holograms display high selectivity, and their reconstruction process results in a convenient conversion of wavelength into angular multiplexing. **(p. 72)**

10:00 am–10:30 am COFFEE BREAK

MESQUITE ROOM

10:30 am–12:10 pm
OWB, PARALLEL SYSTEMS: 1
Bernard H. Soffer, *Hughes Research Laboratory, Presider*

10:30 am (Invited)
OWB1 Smart optical interconnects for high-speed photonic computing, Peter S. Guilfoyle, Frederick F. Zeise, John M. Hessenbruch, *OptiComp Corp.* Multi-element, smart GaAs DANE switching arrays are being developed for free-space optical bus applications and MCM data routing. **(p. 78)**

11:00 am
OWB2 Data/knowledge-base processing on a ODVM platform, Richard V. Stone, John M. Hessenbruch, *OptiComp Corp.* An ODVM (optical digital vector matrix) platform can scan 20,000 pages of text per second for multiple search objects in raw text search applications. **(p. 82)**

WEDNESDAY, MARCH 17, 1993—*Continued*

11:20 (Invited)
OWB3 3-D optical memories, Sadik Esener, *University of California-San Diego.* Abstract not available at press time. **(p. 86)**

11:50 am
OWB4 Comparison of wavelength and angle multiplexed holographic memories, Geoffrey W. Burr, Demetri Psaltis, *California Institute of Technology;* Kevin Curtis, *Northrop.* The storage capacity of an angle/wavelength memory is equal to the number of resolvable elements of the angle/wavelength scanner, rather than crosstalk limited. **(p. 87)**

12:10 pm–2:00 pm LUNCH BREAK

MESQUITE ROOM

2:00 pm–3:30 pm
OWC, OPTICAL NEURAL NETWORKS
Kelvin H. Wagner, *University of Colorado, Presider*

2:00 pm (Invited)
OWC1 Holographic neural networks: a systems perspective, Yuri Owechko, *Hughes Research Laboratories.* A variety of working holographic neural networks have been recently demonstrated in several laboratories, including our own. My talk reviews recent developments, primarily from a system or user point-of-view. **(p. 92)**

2:30 pm
OWC2 Second-order neural network implementation using asymmetric Fabry–Perot modulators, Andrew Jennings, Brian Kelly, John Hegarty, Paul Horan, *Trinity College, Ireland.* A second-order neural network algorithm has been implemented optically, with asymmetric Fabry-Perot modulator quantum well device arrays functioning as optical input devices and weighted interconnects. **(p. 96)**

2:50 pm
OWC3 Content addressable network implementations, Stephen A. Brodsky, Clark C. Guest, *UC–San Diego.* Optoelectronic implementations of three content addressable network (CAN) learning algorithms employing optical computations for supervised learning and recall in supervised, self-organized, and tutored CAN are presented. **(p. 100)**

3:10 pm
OWC4 Random interconnections with ground glass for optical TAG neural networks, Hyuek-Jae Lee, Soo-Young Lee, Sang-Yung Shin, *Korea Advanced Institute of Science and Technology, Korea.* Ground glass provides high-density random interconnections for large-scale optical implementation of neural networks, while SLM is used for adaptive learning of local interconnections. **(p. 104)**

3:30 pm–4:00 pm COFFEE BREAK

MESQUITE ROOM

4:00 pm–6:10 pm
OWD, ANALOG OPTICAL PROCESSING
Soo Young Lee, *Korea Advanced Institute of Science & Technology, Korea, Presider*

4:00 pm (Invited)
OWD1 Optical processing in photorefractive crystals, John H. Hong, *Rockwell International Science Center.* We describe the use of photorefractive crystals in several rf signal processing applications where the unique capability of dynamic holography operating in conjunction with acousto-optic devices provides effective processing systems. **(p. 110)**

4:30 pm
OWD2 Photorefractive phased-array-radar processor dynamics, Robert T. Weverka, Anthony W. Sarto, Kelvin H. Wagner, *Univ. Colorado, Boulder.* We derive, and experimentally verify the dynamic and steady state behavior of a high-bandwidth, large degree-of-freedom phased-array-radar optical processor. **(p. 111)**

4:50 pm
OWD3 Performance evaluation of an acousto-optic wide band correlator system, R. D. Griffin, J. N. Lee, *U.S. Naval Research Laboratory.* We report the performance of an acousto-optic correlator that performs in an integrated system 20–70 times faster than a VAXVector 6410, with potential for a 100X speedup. **(p. 115)**

5:10 pm
OWD4 Optical matrix multiplication using grating degeneracy in photorefractive media, Claire Gu, *The Pennsylvania State Univ.;* Scott Campbell, Pochi Yeh, *UC-Santa Barbara.* We propose and demonstrate a novel method which utilizes grating degeneracy in photorefractive media and an incoherent laser array to implement parallel optical matrix-matrix multiplication. **(p. 119)**

5:30 pm
OWD5 Optoelectronic fuzzy logic inference system using beam scanning laser diodes, Hideo Itoh, Masanobu Watanabe, Seiji Mukai, Hiroyoshi Yajima, *Electrotechnical Laboratory, Japan.* A high-speed optoelectronic fuzzy controller using beam-scanning laser diodes is proposed. The controller uses algebraic product-sum-gravity method with Gaussian membership functions, which is suitable for optoelectronic implementations. **(p. 123)**

5:50 pm
OWD6 Fourier transforms of fractional order and their optical interpretation, David Mendlovic, *Tel-Aviv Univ., Israel;* Haldun M. Ozaktas, *Bilkent Univ., Turkey;* Adolf W. Lohmann, *Angewandte Optik, Germany.* Fourier transforms of fractional order are defined. An optical interpretation is provided in terms of quadratic graded index media and discussed from both wave and ray view points. **(p. 127)**

GRAND BALLROOM CENTER

8:00 pm–9:30 pm
OWE, OPTICAL COMPUTING POSTER SESSION: 2

OWE1 Optical transpose interconnection system, Philippe J. Marchand, Gary C. Marsden, Sadik C. Esener, *UC-San Diego.* The optical transpose interconnection system is a compact, efficient and simple optical interconnection that supports *k*-shuffle, mesh-of-tree, or hypercube optoelectronic architectures. **(p. 132)**

WEDNESDAY, MARCH 17, 1993—*Continued*

OWE2 Theoretical study of the operating conditions for S-SEEDs within cascaded digital optical arrays, Marc P. Y. Desmulliez, John F. Snowdon, Brian S. Wherrett, *Heriot-Watt Univ., U.K.* Tolerances of S-SEED logic gates, within optical circuits, to inter-gate leakage and 2-D nonuniformity are quantified, employing an analytic model developed for the device dynamics. **(p. 136)**

OWE3 Reversal input superposing technique for all-optical neural networks, Yoshio Hayasaki, Ichiro Tohyama, Toyohiko Yatagai, Masahiko Mori, Satoshi Ishihara, *Univ. Tsukuba, Japan.* A new technique using reversal input superposition for optical neural networks is present. Negative values and subtractions are not required in optical implementation with this technique. **(p. 140)**

OWE4 Tandem D-STOP architecture for error backpropagation networks, Gary C. Marsden, Ashok V. Krishnamoorthy, Sadik C. Esener, *UC-San Diego;* Jean Merckle, *Univ. de Haute-Alsace, France.* We present an optoelectronic neural architecture which allows fully parallel on-line learning based on the error backpropagation algorithm. **(p. 144)**

OWE5 Direct optical implementation of lateral inhibition and excitation, Paul Horan, *Trinity College, Ireland.* An optical scheme for implementing lateral excitatory and inhibitory interconnections in a neural network is presented, along with a proof-of-principal simulation. **(p. 148)**

OWE6 Neural algorithms that exploit optical hardware characteristics, John F. Snowdon, *Heriot-Watt Univ., U.K.* The use of optical device characteristics in the formulation of neural algorithms is considered, and the capabilities of resulting systems are analyzed. **(p. 152)**

OWE7 Optical processing system for industrial process control application, Sushila Singh, A. D. Shaligram, *Univ. Poona, India.* This paper reports designing of an optical processor for on-off and continuous types of control. It makes use of optical magnitude comparator based on shadowcasting approach. **(p. 156)**

OWE8 Optical algorithm for a plane image reconstruction from its boundary, Y. B. Karasik, *Tel Aviv Univ, Israel.* The technique of optical computing is applied to develop a new filling algorithm which fills the interior of a plane image in constant time. **(p. 160)**

OWE9 Optical algorithms in geometry, Y. B. Karasik, M. Sharir, *Tel Aviv Univ., Israel.* We present optical algorithms—sequences of optical operations—that solve efficiently many basic geometric problems, motivated by (and applicable in) graphics, pattern matching, etc. **(p. 164)**

OWE10 Algorithm for implementing fully connected optical interconnection networks with broadcast capability, Ahmed Louri, *Univ. Arizona.* We present a new algorithm for implementing fully connected optical interconnection networks with broadcast capability. The algorithm achieves full connectivity between nodes arranged on a two-dimensional plane in constant time. **(p. 168)**

OWE11 Optical data filter for selection and projection, P. A. Mitkas, S. A. Feld, L. J. Irakliotis, C. W. Wilmsen, *Colorado State Univ.* An optical data filter is presented capable of performing selection and projection operations in a relational database environment. The filter is based on arrays of light amplifying optical switch optical gates. **(p. 172)**

OWE12 Polarization- and birefringence-based optical system for reconfigurable weighed interconnects and image shifting, Werner Peiffer, Hugo Thienpont, Michiel Pelt, Roger Vounckx, Irina Veretennicoff, *Vrije Univ. Brussel, Belgium.* A computer driven optical system is presented for dynamically reconfigurable and weighed local interconnects. Videotaped demonstrations and quantitative measurements illustrate actual performances. Future prospects are discussed. **(p. 176)**

WEDNESDAY, MARCH 17, 1993—*Continued*

OWE13 Optical interconnection slab-based static dataflow computer, Alastair D. McAulay, *Lehigh Univ.* A static dataflow computer is proposed and simulated that uses a glass slab as an interconnection network for flip chip bonded processors and memory cells. **(p. 180)**

OWE14 Board-to-board high-speed optical interconnections, J. Jiang, *Duisburg Univ., Germany.* A new concept using a cylinder mirror and particularly formed light-guiding plates is developed for board-to-board high-speed optical interconnections. **(p. 184)**

OWE15 Synchronizing and controlling fast digital optical processors, Vincent P. Heuring, Valentin N. Morozov, *Univ. Colorado at Boulder.* We present solutions to the problems of timing resynchronizing and controlling high-speed optical processors in the context of the gate and strobe paradigm and the more recent time-of-flight paradigm. **(p. 188)**

OWE16 Model of lossless bus structure using erbium fiber amplifiers pumped near 820 nm, Manoj M. Bidnurkar, Steven P. Levitan, Rami G. Melhem, Donald M. Chiarulli, *Univ. Pittsburgh.* A model and simulation results for a lossless tapped fiber bus are presented. Optical amplification is provided by an erbium fiber amplifier using a pump wavelength near 820 nm. **(p. 192)**

OWE17 Simulated annealing applied to placement of processing elements in optoelectronic multichip modules interconnected by computer-generated holograms, D. Zaleta, J. Fan, S. H. Lee, C. K. Cheng, *UC–San Diego.* Results of applying the simulated annealing algorithm to the placement of processing elements based on CGH fabrication limits for optoelectronic multichip modules is presented. **(p. 196)**

OWE18 One-step modified signed-digit addition/subtraction based on redundant bit representation, Hongxin Huang, Masahide Itoh, Toyohiko Yatagai, *Univ. Tsukuba, Japan.* One-step modified signed-digit addition and subtraction, which employ 34 minterms, have been realized with two full parallel optical configurations based on redundant bit representation. **(p. 200)**

OWE19 Concepts for fully integrated FET-SEED circuits and their applications for optical computing, A. J. Wachlowski, D. Rhein, *Alcatel Austria Elin Research Centre, Austria;* F. A. P. Tooley, *Heriot-Watt Univ., U.K.* We describe concepts for computing and sorting applications using FET–SEED technology to achieve higher logic integration. Computer simulation is used extensively to verify the performance. **(p. 204)**

OWE20 Strained multiple quantum well all-optical bistable microcavity operating at 980 nm, J. L. Oudar, D. Pellat, R. Kuszelewicz, R. Azoulay, *France Telecom, France.* Strained InGaAs/GaAs, quantum wells located at the standing-wave antinodes of a high finessse AlAs/GaAs microcavity display all-optical bistability at 980 nm with 10-$\mu W/\mu m^2$ threshold. **(p. 208)**

THURSDAY, MARCH 18, 1993

GRAND BALLROOM WEST

8:30 am–10:00 am
OThA, SMART PIXELS: 1
Richard A. Linke, *NEC Research Institute Inc., Presider*

8:30 am (Invited)
OThA1 Smart pixel optical computing architectures, A. A. Sawchuk, L. Cheng, *Univ. Southern California;* S. R. Forrest, P. R. Prucnal, *Princeton Univ.* Optoelectronic integrated circuit smart pixel arrays implement programmable amplifiers, inverters, logic, bistable switch or latch functions in reconfigurable cellular hypercube processors, shuffle, or sorting networks. **(p. 214)**

9:00 am
OThA2 Field effect transistor–self electro-optic effect device (FET–SEED) circuits for optoelectronic data processing systems, L. M. F. Chirovsky, L. A. D'Asaro, E. J. Laskowski, S. S. Pei, M. T. Asom, M. W. Focht, J. M. Freund, G. D. Guth, R. E. Leibenguth, G. Livescu, R. A. Morgan, T. Mullaly, L. E. Smith, A. L. Lentine, G. D. Boyd, T. K. Woodward, *AT&T Bell Laboratories.* With the monolithic integration of FETs and SEEDs, electronic and photonic technologies merge to enhance each others capabilities, creating the possibility for a whole new range of data processing architectures. **(p. 218)**

9:20 am
OThA3 Cascaded operation of two 128-pixel FET–SEED smart pixel arrays, F. B. McCormick, A. L. Lentine, L. M. F. Chirovsky, *AT&T Bell Laboratories.* We demonstrate cascaded operation of two single transistor FET–SEED smart pixel arrays at 28 Mb/s (14 MHz square wave), limited by optical signal nonuniformities. **(p. 222)**

9:40 am
OThA4 Design and construction of looped parallel processors, S. Wakelin, F. A. P. Tooley, *Heriot-Watt Univ., UK.* Investigation into generic issues in free-space digital optical interconnection, by the construction of looped parallel processors using devices based on self-electro-optic effect device technology is described. **(p. 226)**

10:00 am–10:30 am COFFEE BREAK

GRAND BALLROOM WEST

10:30 am–12:00 pm
OThB, SMART PIXELS: 2
Joseph W. Goodman, *Stanford University, Presider*

10:30 am (Invited)
OThB1 Applications of smart pixels, T. J. Cloonan, *AT&T Bell Laboratories.* Abstract not available at press time. **(p. 232)**

11:00 am
OThB2 Application of self-similar patterns to optoelectronic shuffle/exchange network design, Michael W. Haney, *BDM International, Inc.* A new smart pixel array layout methodology, based on self-similar grid patterns, is proposed to improve packaging efficiency in free-space shuffle/exchange networks. **(p. 233)**

11:20 am
OThB3 Nonblocking optical interconnection networks using electro-optical matrix switches based on VSTEPs, S. Araki, S. Kawai, H. Kurita, K. Kubota, T. Numai, K. Kasahara, *NEC Corp. Japan.* The number of input channels in proposed novel electrooptical matrix switches was evaluated on the basis of their optics. Elemental functions for the switches were certified by using VSTEP and microlens arrays. **(p. 237)**

11:40 am
OThB4 Packaging considerations for planar optical systems, Bruno W. Acklin, Jürgen Jahns, *AT&T Bell Laboratories.* We present a specific example for building an integrated optoelectronic system using the concept of planar optics. Based on issues such as the optical design, thermal management, and fabrication constraints, we discuss various aspects of packaging such a system. **(p. 241)**

12:00 pm–2:00 pm LUNCH BREAK

GRAND BALLROOM WEST

2:00 pm–3:30 pm
OThC, OPTICAL INTERCONNECTIONS: 1
Mark Bendett, *IMRA America, Presider*

2:00 pm (Invited)
OThC1 Free-space optical switching modules, Masayasu Yamaguchi, Ken-ichi Yukimatsu, *NTT Communication Switching Laboratories, Japan.* This paper presents hardware of recent analog and digital free-space photonic switches based on liquid crystal and semiconductor array devices. Their applications are also discussed. **(p. 246)**

2:30 pm
OThC2 Design issues for beam array generation gratings, Rick L. Morrison, Sonya L. Walker, *AT&T Bell Laboratories.* The increasing size and complexity of free-space photonic systems is placing new demands on the design, fabrication, and operation of phase gratings. **(p. 250)**

2:50 pm
OThC3 Packaged optical interconnection system based on photorefractive correlation, Hidemi Takahashi, David Zaleta, Jian Ma, Joseph E. Ford, Yeshayahu Fainman, Sing H. Lee, *UC-San Diego.* We demonstrated a packaged free-space optical interconnection by photorefractive correlation, consisting of a phase code, a $LiNbO_3$ crystal, two CGH lenses, and a detector on a glass substrate. **(p. 254)**

3:10 pm
OThC4 Polarization-selective computer-generated holograms for optical multistage interconnection networks, Joseph Ford, Fang Xu, Ashok Krishnamoorthy, Kristopher Urquhart, Yeshayahu Fainman, *UC-San Diego.* Computer-generated holograms with independent response to two orthogonal linear polarizations were made using etched birefringent substrates. We demonstrate a self-routing binary switch for multistage optical interconnection networks. **(p. 258)**

3:30 pm–4:00 pm COFFEE BREAK

GRAND BALLROOM WEST

4:00 pm–6:10 pm
OThD, PARALLEL SYSTEMS: 2
Jacek Chrostowski, *National Research Council, Canada, Presider*

4:00 pm (Invited)
OThD1 Optically interconnected multichip modules using computer generated holograms, Michael R. Feldman, James E. Morris, John Childers, Mouna Nakkar, Fouad Kiamilev, *Univ. North Carolina at Charlotte.* Progress is proceeding toward the development of an optically interconnected multichip module. The module contains silicon chips with integrated photodetectors. Holograms are used to route optical signals. **(p. 264)**

THURSDAY, MARCH 18, 1993—*Continued*

4:30 pm
OThD2 Massively parallel computers using holographic matrix as a photonic backplane, Michel Charrier, Bruno Houssay, Thierry Lemoine, Sylvain Paineau, *Thomson CSF, France.* Optical interconnect opens new ways for the realization of broadcast communication networks, useful in future generations of massively parallel computers. **(p. 268)**

4:50 pm
OThD3 Massively parallel processing system with an architecture for optoelectronic computing, Masatoshi Ishikawa, Akira Morita, Nobuo Takayanagi, *Univ. Tokyo, Japan.* A new architecture for optical computing is proposed, and an experimental optoelectronic massively parallel processing system as a scale up model of OEIC is described. **(p. 272)**

5:10 pm
OThD4 Multiprocessor system using interboard free-space optical interconnects: COSINE-III, Toshikazu Sakano, Kazuhiro Noguchi, Takao Matsumoto, *NTT Transmission Systems Laboratories, Japan.* COSINE-III is built and tested; it uses 48 bidirectional interboard free-space optical interconnects to three-dimensionally connect 64 processing units. **(p. 276)**

5:30 pm
OThD5 Optically connected parallel machine, D. M. Monro, J. A. Dallas, J. A. Nicholls, *Univ. Bath, U.K.;* M. D. Cripps, *City Univ., U.K.;* W. A. Crossland, *Cambridge Univ., U.K.* A multiprocessor array achieves 640 Mbit/sec/channel with free space optics switched by a 10-μsec full crossbar 64 × 64 liquid crystal SLM, with potential for improvement. **(p. 280)**

5:50 pm
OThD6 Cellular processing with diffractive optical elements, Andrew Kirk, Tomohira Tabata, Masatoshi Ishikawa, *Univ. Tokyo, Japan.* A high-speed optoelectronic cellular processing system, which employs a reconfigurable diffractive element to perform a shift-invariant interconnection operation, is described and experimental results are presented. **(p. 284)**

GRAND BALLROOM WEST

8:00 pm
OPDP, OPTICAL COMPUTING POSTDEADLINE SESSION
John E. Midwinter, *University College London, U.K., Presider*

FRIDAY, MARCH 19, 1993

GRAND BALLROOM WEST

8:30 am–10:00 am
OFA, OPTICAL INTERCONNECTIONS: 2
Phillip J. Anthony, *AT&T Bell Laboratories, Presider*

8:30 am (Invited)
OFA1 GaAs heteroepitaxial growth on submicron CMOS silicon substrates, S. K. Tewksbury, L. A. Hornak, H. Nariman, S. McGinnis, *West Virginia Univ.* Preliminary results of our study of the CMOS compatibility of GaAs heteroepitaxial growth on wafers containing twin-tub V CMOS devices (0.9 micron drawn features) are reviewed. **(p. 290)**

9:00 am
OFA2 Four-foci slanted axis Fresnel lens for planar optical perfect shuffle, C. D. Carey, D. R. Selviah, J. E. Midwinter, *Univ. College London, U.K.*, S. H. Song, E. H. Lee, *Electronics and Telecommunications Research Institute, Korea.* A binary, reflective, amplitude, four-foci, slanted axis, computer generated holographic Fresnel lens performs a white light perfect shuffle in an imaging planar optic configuration. **(p. 291)**

9:20 am
OFA3 Experimental implementation of a 25-channel free-space optical switching architecture, Andrew Kirk, *Univ. Tokyo, Japan*, Stuart Jamieson, Hussain Imam, Trevor Hall, *King's College London, U.K.* A matrix-matrix multiplier crossbar switch is described which is based on the holographic multiple imaging of a Gaussian beamlet array. Experimental results are presented. **(p. 295)**

9:40 am
OFA4 Bandwidth as a virtual resource in reconfigurable optical interconnections, Donald M. Chiarulli, Steven P. Levitan, Rami G. Melhem, Chunming Qiao, *Univ. Pittsburgh.* Locality-based control paradigms which are directly analogous to virtual memory management are adapted for message routing and control of reconfigurable optical interconnection networks. **(p. 299)**

10:00 am–10:30 am COFFEE BREAK

GRAND BALLROOM WEST

10:30 am–12:00 pm
OFB, DIGITAL OPTICAL COMPUTING
Frank A. Tooley, *Heriot-Watt University, U.K., Presider*

10:30 am (Invited)
OFB1 Digital optical computing, K. H. Brenner, *Univ. Erlangen, Germany.* Abstract not available at press time. **(p. 304)**

11:00 am
OFB2 Spatial soliton dragging gates and light bullets, Kelvin Wagner, Robert McLeod, *Univ. Colorado.* Spatial soliton dragging provides a low-latency, cascadable, three-terminal optical logic gate with gain that can be extended to three-dimensions using light bullets. **(p. 305)**

11:20 am
OFB3 Case for all-optical digital computing, Miles Murdocca, *Rutgers Univ.* Reconfigurable gate-level optical interconnects can provide a performance advantage over an opto-electronic approach, which may compensate for the increased expense of an all-optical approach. **(p. 309)**

FRIDAY, MARCH 19, 1993—*Continued*

11:40 am
OFB4 Optical implementations of the modified signed-digit algorithm, Shaomin Zhou, Scott Campbell, Pochi Yeh, *UC-Santa Barbara;* Hua-Kuang Liu, *Jet Propulsion Laboratory.* We propose and demonstrate novel optical computing architecture utilize spatial and spectral coding for the implementations of the modified signed-digit algorithm. **(p. 313)**

12:00 pm–2:00 pm LUNCH BREAK

GRAND BALLROOM WEST

2:00 pm–3:10 pm
OFC, BIT SERIAL OPTICAL COMPUTING
Matthew W. Derstine, *Optivision, Presider*

2:00 pm (Invited)
OFC1 An integrated optics, stored program computer, Harry F. Jordan, Alan R. Mickelson, Bart Van Zeghbroeck, I. Januar, *Univ. Colorado.* A serial stored program, digital optical computer has been implemented and operated. Parameters of an integrated version are derived from current integrated optoelectronics research results. **(p. 318)**

2:30 pm
OFC2 Implementation of a general purpose stored program digital optical computer, P. T. Main, R. J. Feuerstein, V. P. Heuring, H. F. Jordan, J. R. Feehrer, C. E. Love, *Univ. Colorado.* The first general purpose stored program optoelectronic computer with optical signals flowing between gates has been constructed. Its construction and operation are described. **(p. 322)**

2:50 pm
OFC3 Time of flight packet synchronizers, Carl E. Love, Harry F. Jordan, *Univ. Colorado.* High-speed optical packet communication networks need to synchronize data from distinct sources. Two architectures for flip-flop free, delay line based synchronizers are demonstrated. **(p. 326)**

3:10 pm–3:20 pm
CLOSING REMARKS
H. S. Hinton, *AT&T Bell Laboratories*

Joint Optical Computing/ Photonics in Switching/ Spatial Light Modulators Plenary Session

JTuC 2:00pm-5:30pm
Grand Ballroom East

B. Keith Jenkins, *Presider*
University of Southern California

Joseph W. Goodman, *Presider*
Stanford University

Extended Generalized Shuffle Networks

G. W. Richards

AT&T Bell Laboratories
Room 2E-252
263 Shuman Blvd.
Naperville, Il. 60566-7050

Tel. (708) 713-5425
Fax. (708) 713-7951

1.0 Introduction

A successful network architecture generally must consider the technology being proposed for implementation. A design suitable for an electronic application may prove to be unwieldly or even impossible to implement in the photonic domain. Basically this is due to the limitations and constraints imposed by the optical technology at hand.

The talk provides an overview of a new class of networks called Extended Generalized Shuffle (EGS) networks[1],[2] and considers how various attributes of these networks can address some photonic switching constraints. From a purely mathematical and topological viewpoint EGS networks are technology independent and they will be handled as such in subsequent more detailed papers[3],[4]. However, our current focus is on photonic switching and we will thus limit most of our discussion to those aspects of EGS networks that primarily impact the implementation of photonic switching networks. The talk is primarily intended to bridge some of the gap between photonic switching technology and switching network theory.

The following two sections catalog some constraints imposed by free-space and guided-wave photonics, respectively. The fourth section briefly summarizes how EGS networks can deal with such contraints.

2.0 Some Free-Space Photonic Switching Constraints

Free-space photonic network implementation assumes the use of symmetric self electro-optic devices (S-SEEDs)[5] as logic gates in the switching modules of the network.

2.1 Switching Module Fan-Out/Fan-In Limitations

There are several reasons for preferring small values of fan-out and fan-in when dealing with free-space optics[6]. If the fan-out in a particular system is large, then the optical power that is emitted from a single S-SEED must be divided before being routed to the many detecting S-SEEDs, and the optical power arriving at any one of the detecting S-SEEDs will be relatively low. Since the maximum switching speed of an S-SEED is directly proportional to the amount of optical power that sets the device, smaller fan-outs will yield faster switching speeds.

Smaller values of fan-in are also desirable within a system because the signal-to-noise ratio at the input of any S-SEED will increase as more signals are fanned into the device. The bit error rate of the system will also increase. In addition, larger values of fan-out and fan-in will typically require more complicated beam-steering optics which tend to increase the overall system cost.

Reference [6] reports a lossless beam splitting and recombination technique that suggests advantages for limiting fan-out and fan-in to values of two between stages of logic. This in turn suggests the use of switching modules having no more than two inputs and two outputs. Such modules may or may not be conventional crossbar switches; an issue we consider next.

2.2 Switching Module Functionality

The functionality provided in switching modules should consider overall network implementation complexity and efficiency. For example, consider two switching modules A and B and suppose that more B modules than A modules are required to construct a non-blocking NxN network. We next consider the implementation complexity (no. of S-

SEEDs, no. of stages, no. of control beams, etc.) of the two module types. If the complexity of module B is less than that of module A, we may find that a network using B modules has less overall complexity than one using A modules (in spite of the fact that more B modules than A modules are required to construct the network).

2.3 Switching Stage Uniformity

One important advantage in having uniform or identical switching stages is that only one type of switching stage needs to be fabricated. However, such uniformity requires that each switching stage has equal numbers of inputs and outputs (because to fully interconnect all of the outputs of a given stage with all of the inputs of the next stage requires that these numbers of inputs and outputs are equal and hence, via uniformity, that the numbers of inputs and outputs on all stages are equal). Thus, switching stage uniformity implies that the number of interconnections between stages is the same throughout the network. We consider this further in the next section.

Efficient, high performance EGS networks can be designed with uniform switching stages. In fact this constraint carries with it almost no disadvantages.

2.4 Interstage Interconnection

The "free-space" descriptor for photonic switching networks refers to the means of interconnecting successive stages of switching modules. It is advantageous for this technique to have constant numbers of interconnections between stages (as mentioned above). Two additional interconnection attributes, which have been found to be particularly convenient when employing free-space optics, are symmetry and stage-to-stage pattern invariance.

The so called "crossover"[7] interconnection pattern has constant numbers of interconnections between stages and exhibits symmetry, but does not result in stage-to-stage invariance. However, the stage dependent variations are such that they can be achieved by selecting an appropriate prismatic mirror array for each stage. We are willing to accept these stage-to-stage variations because the crossover interconnection pattern can be shown to yield EGS networks.

2.5 Switching Stage Size and Number of Stages

There are usually advantages in keeping both the switching stage size and the number of stages small. These advantages relate to such aspects as efficiency, reliability, and power. For example, as the switching stage size is increased, the size of the S-SEED array that implements the logic of the switching stage must also increase. As a result, the optical components (lenses, beam-splitters, etc.) within a stage must be capable of providing diffraction-limited imaging over larger fields of view. In addition, system lasers must provide more optical power if larger S-SEED arrays are used. Both of these requirements will increase the cost of a single stage within the system.

If, on the other hand, the number of stages is increased, then the overall system cost will begin to increase. Also, the expected availability of the system will decrease as the number of stages is increased, because the components within the multiple stages have non-zero failure rates. Unfortunately, it is usually not possible to have small values simultaneously for both switching stage size and number of stages. What sort of compromises are possible?

S-SEED photonic EGS switching networks generally exhibit the following helpful attribute. For a given number of input and output terminals and for a given probability of blocking (including zero), as the S-SEED array size increases (decreases), the required number of stages of S-SEED arrays tends to decrease (increase). Thus, these networks give the designer the capability to trade-off between switching-stage size and number of stages.

3.0 Some Guided-Wave Photonic Switching Constraints

Guided-wave photonic network implementation assumes the use of lithium niobate couplers[8] as the switching modules of the network.

3.1 2x2 Basic Element

The lithium niobate coupler is inherently a 2x2 device. Thus, we assume the use of switching modules having no more than two inputs and two outputs. As with free-space photonics, such modules may or may not be viewed as conventional crossbar switches.

3.2 Switching Module Functionality

Crosstalk due to imperfect lithium niobate couplers can grow to an undesirable level in a large network. One solution to this problem is to allow no more than one active signal in any coupler[9]. This means that a coupler may be unavailable for use even though the desired input and output of the coupler are both idle. EGS networks are able to deal directly with this constraint in network design.

3.3 Interstage Interconnection

To keep losses at a minimum, the waveguide bends between stages of couplers have relatively large radii of curvature. Additionally, the couplers themselves are relatively large in comparison to the substrate wafer used in the fabrication process. The result is that there is low coupler density and therefore limited switching functionality per substrate module. The challenge for EGS networks is to be able to utilize such low functionality modules in the design of efficient high functionality networks.

3.4 Switching Stage Size and Number of Stages

As with free-space networks, there are usually advantages in keeping both the switching stage size and the number of stages small. Overall crosstalk and loss are two items of consideration. EGS networks allow trade-offs between switching stage size and the number of stages. As one increases (decreases) the other decreases (increases). Furthermore, in the most efficient configurations, the number of stages increases slowly (log N) in comparison to N (the number of inlets and outlets).

4.0 EGS Network Attributes

The following subsections catalog a few of the EGS network attributes that are helpful in dealing with photonic switching constraints.

4.1 Global Generalized Conditions for Non-Blocking Operation

It is possible to establish conditons for non-blocking EGS networks in very general terms. These conditions do not impose any constraints on the number of inlets or outlets in a network or any relationship between these two values. There is no symmetry required in the network. There are no constraints on the number of stages in the network. There is no relationship required between the size of the switching modules in one stage and any other stage. Also, the size of the switching modules in any particular stage are not generally constrained.

The major impact here is that for the most part switching networks can be designed with arbitrarily sized modules and an arbitrary number of stages. Thus, 2x2 modules can be utilized easily and switching stage uniformity becomes a natural outgrowth.

4.2 Multiple Stage Modularity

Under certain easily met conditions EGS networks exhibit the following attribute. For specifc (not necessarily consecutive) stages i and j, there exist subsets of switches in both of these stages such that every switch in the stage-i subset has a path to every switch in the stage-j subset and paths to no other switches in stage j. This allows networks to be constructed via multi-stage modules, each of which is entirely unconnected from other modules in the corresponding stages.

4.3 Different Types of Switching Module Functionality

The analysis of EGS networks allows switching modules to have functionalities less than that of conventional crossbar switches. Such reduced functionality usually relates to the sizes of specified subsets of inlets and outlets on the module that can be individually connected to each other. For example, one type of module may only be able to connect all inlets to all outlets or no inlets to no outlets. Efficient EGS networks can be constructed from such modules.

4.4 Network Isomorphisms

Two networks may have the same connectivity, differing only in the way their switching modules and links are labeled or in the way they are drawn. Any two such networks are said to be isomorphic. This relationship divides the collection of all networks into isomorphic classes. It is the fortunate case that many isomorphic classes of EGS networks have members that are amenable to photonic implementation. Formal mapping functions provide the ability to similarly analyze and control many of the members of a given isomorphism class, thereby allowing technology considerations to dictate the particular member chosen for implementation.

4.5 Network Shape Trade-Offs

The generality inherent in EGS networks allows a network with a given number of inlet and outlets to have many different "shapes". More specifically, one can typically trade-off the number of modules per stage with the number of stages. Thus network designs can be tailored to accommodate the needs of various technologies.

4.6 Other Attributes

The theory of EGS networks encompasses other dimensions of switching applications such as routing and control. These items and others have also found value in the implementation of photonic networks. Unfortunately, the limited scope of this summary does not allow further elaboration of such topics.

5.0 References

[1] G. W. Richards, *U.S. Patent Numbers: 4,993,016 and 4,991,168.*

[2] T. J. Cloonan, G. W. Richards, A. L. Lentine, F. B. McCormick, and J. R. Erickson, "Free-space photonic switching architectures based on extended generalized shuffle networks", Appl. Opt. 31, pp. 7471-7492 (1992).

[3] G. W. Richards and F. K. Hwang, "Extended generalized shuffle networks: sufficient conditions for strictly nonblocking operation", in preparation.

[4] G. W. Richards and F. K. Hwang, "Extended generalized shuffle networks and switching module functionality", in preparation.

[5] A. L. Lentine, H. S. Hinton, D. A. B. Miller, J. E. Henry, J. E. Cunningham, and L. M. F. Chirovsky, "Symmetric self-electrooptic effect device: optical set-reset latch, differential logic gate, and differential modulator/detector", IEEE Journ. of Quant. Elect. 25, pp. 1928 - 1936 (1989).

[6] F. B. McCormick and M. E. Prise, "Optical circuitry for free-space interconnections", Appl. Opt. 29, pp. 2013-2018 (1990).

[7] J. Jahns and M. J. Murdocca, "Crossover networks and their optical implementation", Appl. Opt. 27, pp. 3155-3160 (1988).

[8] H. S. Hinton, "Photonic switching using directional couplers", IEEE Communications Magazine, Vol. 25, No. 5, May 1985, pp. 16-26.

[9] K. Padmanabhan and A. N. Netravali, "Dilated networks for photonic switching", IEEE Transactions on Communications, Vol. COM-35, No. 12, pp. 1357-1365, December 1987.

ATM Objectives and Requirements For Next-Generation Networks

Kai Y. Eng
AT&T Bell Laboratories
Room 4F525
101 Crawfords Corner Road
Holmdel, NJ 07733-3030
Phone: (908) 949-2201 Fax: (908) 949-9118
Email: kye@boole.att.com

1. Introduction

Worldwide activities on ATM (Asynchronous Transfer Mode) have been intensifying with rapidly evolving standards (CCITT and ATM Forum) [1,2]. Applications include multimedia services, high-speed LAN's, central-office switches and high-speed digital crossconnects. Potential large-scale network conversion into ATM is being considered and debated within various research and development communities. The push for deploying ATM to upgrade the existing network infrastructure has sometimes been compared to the "analog-to-digital revolution". The primary incentive to do so stems almost entirely from its service flexibility. Here, we discuss various aspects of ATM networking, emphasizing the transport objectives and requirements for next-generation networks. We will begin with the basic notions of ATM networking and then shift to a specific example of a VP (Virtual Path) transport network using an integrated ATM Crossconnect as a key network element.

2. Basics of ATM Networking

In ATM networking, all user-generated information (voice, video and data) is converted into standard fixed-size packets called "cells" for switching and transport. This process of converting the original data into ATM cells is called "adaptation". An ATM cell has a fixed size of 53 octets consisting of a 5-octet header followed by a 48-octet payload (Fig. 1). The header contains primarily the routing information for the cell (VCI and VPI values). Unlike traditional synchronous TDM transmission systems where data are sent via preassigned time slots, ATM traffic is transported by cell multiplexing without dedicated slots, and cells from a given source would appear to the network as asynchronous traffic. A simplified depiction of an ATM network connecting two users, A and B, is illustrated in Fig. 2. In order to initiate a communication connection between A and B, a call set-up procedure has to be invoked, e.g., a call request from A to B via ATM Switch I (signaling and control not shown in Fig. 2). As part of the call set-up process, ATM Switch I needs to inform A's ATM adapter proper VCI and VPI values to use and also to assure that these header parameters are recognized by various network elements along the assigned route for routing. Let us focus on the routing function first.

An ATM end-to-end connection from A to B is designated as a Virtual Circuit and hence a VCI in the cell header. A specific path from ATM Switch I to ATM Switch II, however, is called a Virtual Path which may contain many Virtual Circuits between various source-destination pairs connected to ATM Switches I and II. Along a specific VP, there may be multiple ATM Crossconnects (XC's) which are VP switches. Therefore, a VP is essentially a specific routing path from an originating ATM Switch to a destinating ATM Switch through a series of ATM XC's. All cells from a given source in the same connection traverse the same Virtual Path, and their original sequence is thus maintained. The VPI value in the header designates the VP, but its use requires a translation process explained later in the next section.

The transmission line standard between ATM network elements (Switches and XC's) is assumed to be SDH (or SDH), e.g., 155 Mb/s OC-3c, 622 Mb/s OC-12c or 1.4 Gb/s OC-48. As specified in the standards, ATM cells are mapped into the payload of the SDH envelope. Since the VPI value has 12 bits, each transmission link can carry up to 4096 unique VP's. The processing of the SDH overhead and ATM cells in the payload is an important part of the line interface at each ATM XC (often called "line cards"). The delivery of cells according their VP, namely VP transport, plays an important role in ATM networking because of its flexible service and fast restoration capabilities [3]. We will focus on the operation of VP transport by describing an overview of an integrated ATM XC.

3. ATM Crossconnect Functions and Requirements

When ATM cells are carried as payload in the SDH envelope for optical transmission, they have to be recovered before switching can take place at each ATM XC. The function of receiving the optical signal, terminating its SDH overhead and recovery of its payload (i.e., ATM cells) is usually called the LTE (lightwave terminal) function. The VP switching of cells is of course the main goal of the XC. VP switching here means the routing of the incoming cells to the appropriate output ports according to their VPI values on a cell-by-cell basis. Although not required, today's technology permits an integrated design whereby both the LTE and the VP switching functions are combined in a single machine. The LTE part is implemented mostly in the line cards, and the VP switching in the "XC fabric". For subsequent discussions, we only consider an integrated ATM XC.

Various essential functions required in an integrated ATM XC are summarized in Fig. 3. After optical detection, the received SDH signal is descrambled and frame synchronized so that the overhead bytes can be retrieved. Processing of these overhead bytes is required to provide several maintenance functions such as line integrity, line protection switching, inter-XC data communications, etc. Furthermore, pointer processing has be performed to establish synchronization for recovering the "floating" frames (called virtual containers) of ATM cells in the payload. The beginning of each cell has then to be identified within the virtual container. After the cells are extracted, they have to be properly processed before the actual switching, and this part may either be implemented in the line card or as part of the XC fabric.

In ATM cell processing, OA&M (operation, administration and maintenance) cells are separated from data cells. OA&M cells may be destined for the local system controller or for further transport. The data cells are first processed with HEC (header error correction). Cell headers with 1-bit errors can be corrected, and those with more than 1-bit errors are discarded immediately. The VPI value is then examined in a table look-up for validation and also to identify the output port for the routing of each cell. Cell routing is implemented in the XC fabric which is functionally identical to a conventional ATM switch fabric. As such, cell buffering is required because multiple cells may destine for the same output port at the same time. This aspect has been well understood in ATM switching. In addition to the routing, the most most important processing is perhaps the translation of the incoming VPI value in each cell to a new value (preassigned by network control). That is, for each cell transit through the ATM XC, there is a unique mapping of its incoming VPI value to its outgoing value. Consequently, a VP can be characterized by a unique series of VPI values along the various physical links in its route. Doing so allows for maximum utilization of the VPI values (limited to 12 bits) in each physical link.

At an output port of the XC fabric, cells from different inputs are multiplexed together to form a stream. They have to be individually checked for errors and to verify proper routing (plus internal diagnostics) before passed on as input to a SDH processor. In this SDH processor (output line card), the ATM cells are inserted into the payload envelopes, and new pointer and other overhead bytes are included in the final assembly of the whole SDH signal for optical transmission.

The capability of electronics is advancing rapidly to meet the requirements of high-performance ATM network elements. An integrated prototype ATM XC recently demonstrated includes an 8×8 ATM fabric operating at 2.5 Gb/s and with multiple-rate line cards (OC-3c to OC-48) [4]. This is equivalent to a total system capacity of 20 Gb/s, supporting up to 128 bidirectional OC-3c (155 Mb/s) interfaces or up to eight bidirectional OC-48 (2.4 Gb/s) interfaces. Expansion to much larger fabric sizes is possible and is being researched.

4. Conclusions

VP transport is a fundamental aspect of ATM networking. It provides flexible service and fast restoration capabilities in next-generation networks. An ATM XC capable of cell-by-cell routing is a key network element to support VP transport. In addition, extensive SDH and ATM cell processing is required in implementing the ATM XC, today's VLSI technology is adequate to meet these challenges efficiently and economically.

REFERENCES

[1] Martin de Prycker, *Asynchronous Transfer Mode*, Ellis Horwood, 1992.

[2] *CCITT Recommendations*, I.113, I.121, I.150, I.211, I.311, I.321, I.327, I.361, I.362, I.363, I.413, I.432 and I.610.

[3] K. Sato, H. Ueda, and N. Yoshikai, "The Role of Virtual Path Crossconnection," *IEEE LTS Magazine*, August 1991, pp. 44-54.

[4] K. Y. Eng, M. A. Pashan, R. A. Spanke, M. J. Karol and G. D. Martin, "A High-Performance Prototype 2.5 Gb/s ATM Switch For Broadband Applications," *IEEE Globecom '92 Conference Record*, Vol. 1, Dec. 6-9, 1992, pp. 111-117.

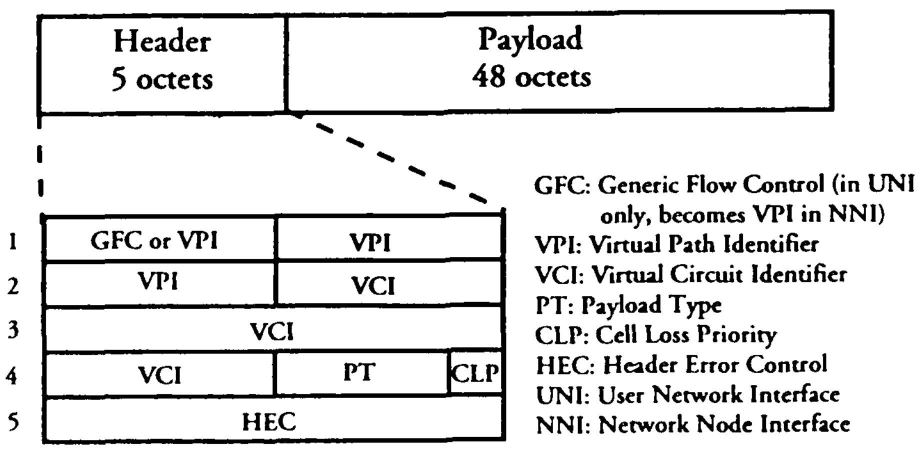


Fig. 1 ATM Cell

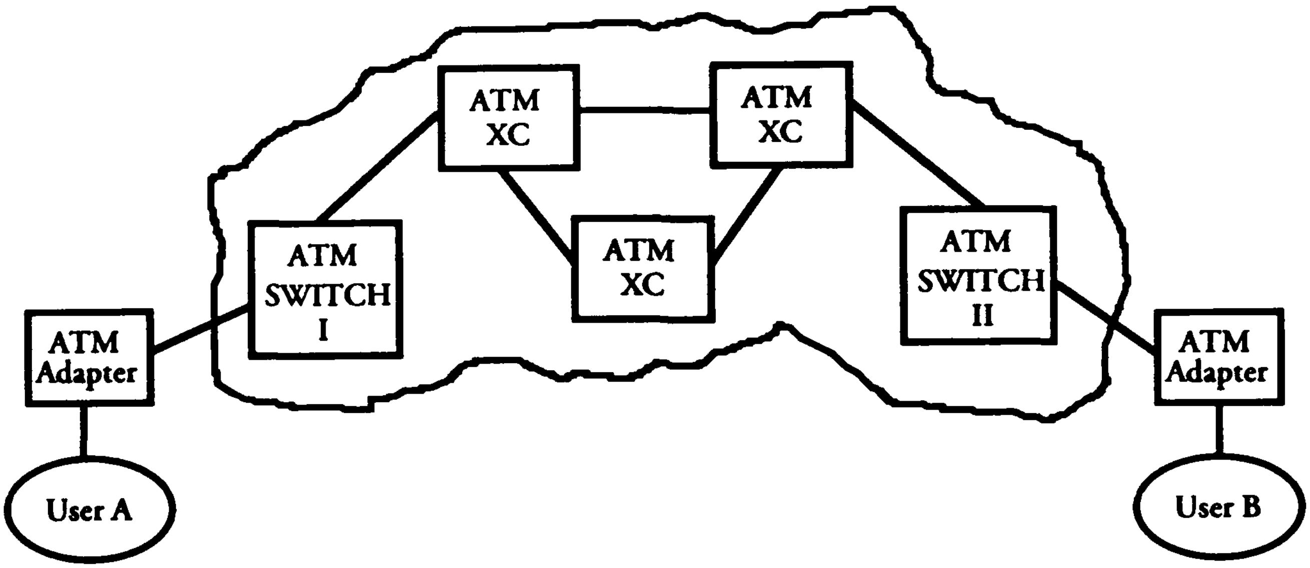
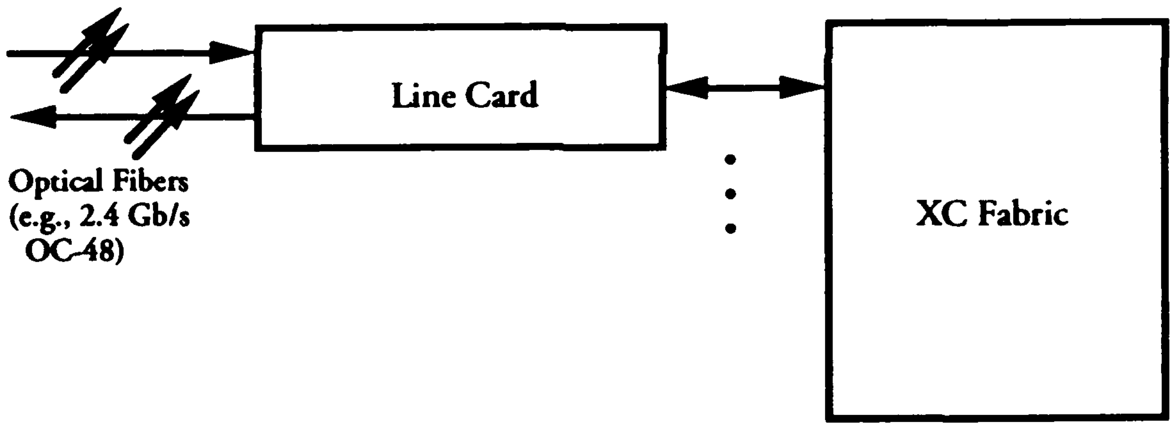


Fig. 2 An ATM Network Example



- SDH Overhead Termination
  - SDH framing
  - Timing recovery and synchronization
  - SDH overhead extraction
  - Pointer processing
  - Protection switching
  - Line and section parity check
  - Cell framing
- ATM Cell Processing
  - Cell synchronization
  - VPI validation
  - VPI table look-up
  - HEC
  - Internal diagnostics
  - OA&M
- XC
  - Physical routing
  - Cell buffering
  - Protection switching

Fig. 3 Essential Functions of An Integrated ATM XC

Photonics in switching: European systems demonstrators and the long term perspective

Lars Thylén *
Dep of Microwave Engineering
Royal Institute of Technology
S-100 44 Stockholm Sweden

January 13, 1993

Status in the area of photonics in switching in Europe, highlighted by systems demonstrators, is reviewed,and the long term perspective for photonics in switching is discussed.

1 INTRODUCTION

It is widely recognized that the progress and role of photonic switching have not paralleled that of fiber optics point to point communications, where new avenues for fiber optics keep opening, most recently in the shape of *fiber to the home* and the associated introduction of broadband services. The discussions concerning fiber to the home, and the corresponding systems issues, seem to have been only marginally influenced by the possibilities offered by photonic switching. Part of the explanation for this state of affairs is the competition from electronics, and the uncertainty of the actual bandwidths required by different services and subscriber categories, but mostly the lack of clear and verifiable systems solutions where photonic switching offers significant and unique advantages. In addition, we have the comparative immaturity of the device technology required as well as the lack of a practical and reasonably standardized way of assembling optical systems. This situation is reflected in the comparatively small number of systems demonstrators and the virtually nonexistent field trials to date. Ref [1], [2] are exceptions here, but they can not be described as fully fledged field trials.

This paper discusses development over the last few years concentrating on the European scene, where several European research programs (COST as well as RACE) have included projects (especially the RACE OSCAR R1033) more or less devoted to photonic switching. Some current representative systems demonstrators are described. Comments on the relationship between electronic and optical switching, based on switching energy, pertinent for ATM switching, are finally made.

This paper concentrates on *guided wave switching systems and devices* , there are strong proponents of *free space 3D systems and technology* , [3], [4]. However, the situation here is even less mature as far as technology is concerned, and systems proposals and technology demonstrated so far are no more convincing than the guided wave concepts. The rationale and motivation for *integrated photonics* in photonic switching is the same as those underlying the unparalleled success of *integrated electronics* .

2 DEVELOPMENTS IN PHOTONICS IN SWITCHING

The rationale for photonic switching is the uncontested bandwidth and loss of the optical guided wave transmission as well as the speed allowed by optical interactions. Different attempts have been made to

*Also with Fiber Optics Research Center, Ericsson Telecom AB, S-126 25 Stockholm, Sweden

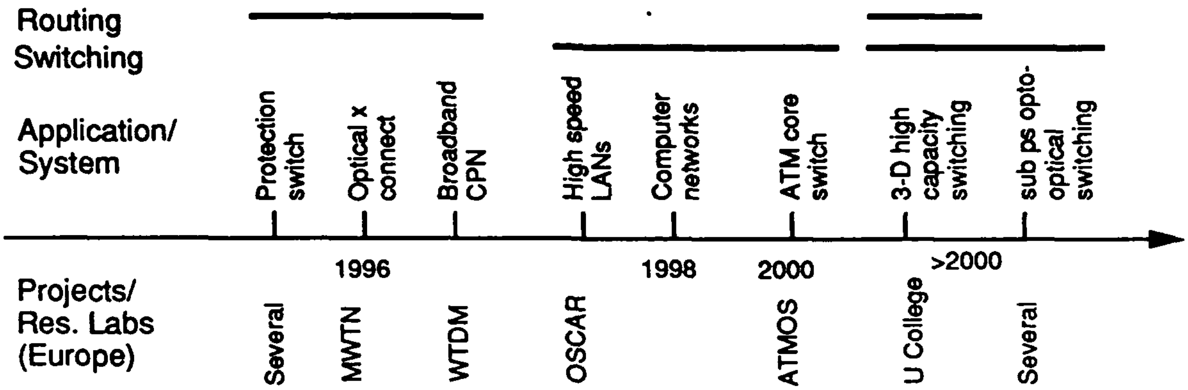


Figure 1: Evolution scenario of photonic switching. Years indicate possible time of field trial. MWTN, WTDM, ATMOS and OSCAR are RACE projects, mentioned in the text

shape this into new systems architectures, ("optical ether"). Broadly speaking, this has led to a number of schools:

A) Circuit type low speed switching, involving switching in space and wavelength

B) STM, ATM and PTM type switching, where optical temporal switching is involved,in some cases in conjunction with wavelength and space division switching.

C) 3D optical interconnects in combination with OEICs, where the processing is (mainly) electronic and the transmission optic.

D) Whereas the above approaches can in essence be classified as *optical interconnect under electronic control*, there also is the possibility for *opto-optical switching* , e g via soliton interactions [5]

Crucial in all attempts to apply temporal switching is the lack of an optical RAM type memory, comparable in integration and performance to electronic ones, since various degrees of synchronization and storage are required. It appears that only wavelength and space division switching make use of the bandwidth in a way commensurate with the bandwidth of the transmission medium, thus creating a novel fiber optic "network" that is to a degree bitrate and coding independent ("transparency"). It should be borne in mind that we are dealing with an *analog, nonlinear* network. However, time division switching type network structures have been suggested and partly demonstrated, where optical switching functions are simple and in some cases memoryless ([6], [7]), again taking advantages of the basic features of the optical transmission medium. See also [8]. One could attempt to structure the development in photonic switching according to Fig 1.

3 SYSTEMS DEMONSTRATORS

The systems demonstrators, which in some sense can be labeled as photonic switching ones, currently developed within the RACE program, are part of the following projects: Wavelength and time division multiplexing (WTDM) broadband CPN network, Multiwavelength transport network (MWTN) and ATM optical switch (ATMOS). The first two are essentially in category A) in section 2, the third in category B). All these projects are based on RACE I projects. In addition there is a German program, which addresses ATM type switching. The talk will discuss these projects, two of which are described below. Fig 2 shows the basic structure of the R2039 ATMOS demonstrator [9]. This utilizes synchronized ATM cells at the input (for all practical purposes of electronic origin). The cell encoder wavelength encodes the packets utilizing wavelength converters (eg in the shape of bistable DBR lasers). Contention is resolved by a cell buffer block with K fiber optic delay lines (K=16 in a suggested system; with no contention, this entire block can of course be deleted). Power splitters and optical gates (semiconductor laser amplifiers) route the packets to the pertinent delay line. The packets (still of course synchronized), enter a space switch in the shape of a star coupler, with optical filtering at the output. Four layers of this type, each comprising a Clos net with 3 stages of 16 × 16 switches will have a total throughput of 10 Tb/s, with a 10^{-9} cell loss rate for 16 cell buffers (fig 2), with $10\mu s$ delay. The line rate is 2.6 Gb/s. A rigorous comparison with an all electronic system still remains to be done and is a challenging research topic. It

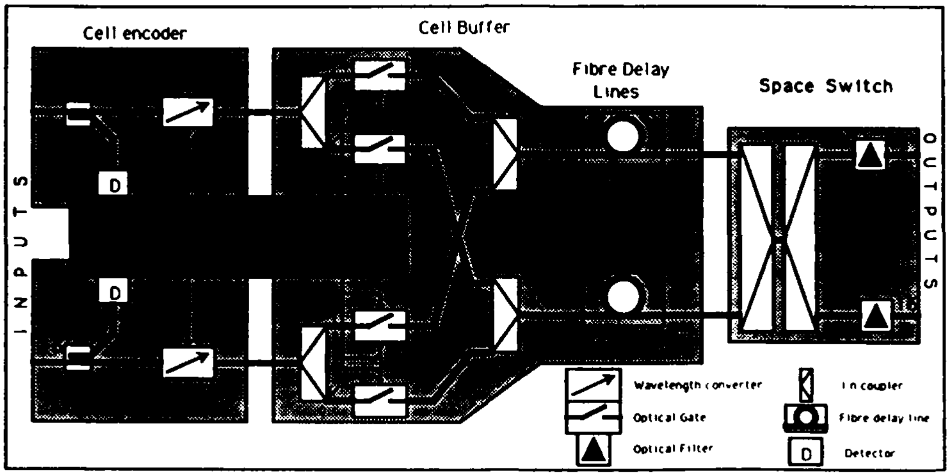


Figure 2: ATM optical switch (ATMOS) system demonstrator

should be noted that impressive results have been reported recently on fiber optic delay lines, applicable to this concept [10]. The ATMOS system presents a number of very challenging device requirements, notably wavelength converters and optical storage. Fig 3, on the other hand, represents an example of a *routing system* [11], see also [12]. The structure in fig 3 performs routing in the space and wavelength domains. Hence, we are not concerned with a reconfiguration of the network, more rapid than that called for by eg protection switching ($> \mu s$). The core of the network is formed by tunable lasers in the transmission system, tunable filters and space switches in the switching system and fiber amplifiers in a line system. Also important (but historically given little attention in systems like this) is the control system. Since the signals are not immediately available in electronic form (with a few exceptions, such as laser amplifiers [13]), the control system has to be structured accordingly, and the devices provided with control interfaces, the need of which is not superficially obvious. Two issues of prime importance are wavelength referencing and power equalization. The system in fig 3 constitutes a logical extension to the existing transport network concepts, adding flexibility, reliability as well as resilience, by performing routing in a frequency and code transparent way. The systems claims made here (the underlying systems rationale is about the same in WTDM) are maybe not as far reaching as for ATMOS, but the application appears reasonably near term. In fact, in view of the continued development of the transport network with electronic cross connects, the optical cross connect appears to be a good candidate for a fairly imminent application, being uncontested by electronics.

4 HIGH SPEED SWITCHING: PHOTONICS OR ELECTRONICS OR BOTH?

A view often advocated in the optics community is to justify photonic switching by extreme speed and superior bandwidth. However, MODFET transistors have been reported with speeds up to 500 GHz, and several roads into the THz realm exist: Superconductors, quantum interference, single electron transfer, and brute force scaling of dimensions. The last exercise has given uninterrupted *exponential* performance growth since the 40s! "Fundamental" limits appear to be no hindrance to continued development for the next few decades. This means that electronic integration levels, switch energy as well as speed will continue to improve. Concerning integration and switch energy, it appears that basic considerations prevents photonics from presenting a viable alternative to electronics [14]. Hence, the combination of photonics and electronics has to be given serious consideration. One example could be low level of photonic integration (where dissipation is less of a concern) for extremely high speed (sub ps) multiplexing and demultiplexing, with the slower and complex (more integration intensive) tasks carried out by electronics.

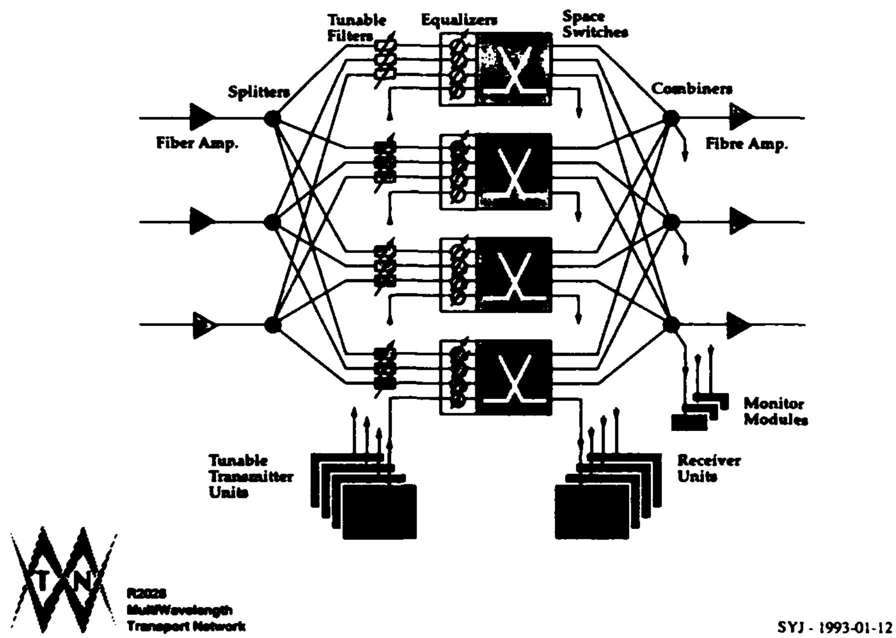


Figure 3: MWTN system demonstrator

5 CONCLUSIONS

This paper has described two representative examples of photonic switching demonstrators, showing different approaches, differing device requirements and addressing different time perspectives. It should be emphasized that photonic switching can not be seen in isolation but only from a total systems perspective. Some research items of photonic switching can be identified: Rigorous research on electronic vs photonic switching, including systems considerations; the potential of extremely high speed, low integration switching using nonlinear optical interactions, such as solitons; 3D reconfigurable interconnect, combined with OEICs. While these are certainly worthwhile to pursue, it appears that the span to the present needs to be bridged by more applied research; thus field trials are highly desired in the next few years, unless photonic switching is to share the fate of at least part of the field of optical computing.

References

[1] D Edinger, P Duthie, G R Prabhakara, "A new answer to fiber protection", Telephony, (1990), p 53

[2] T G Lynch et al, "Experimental field demonstration of a managed multinoded reconfigurable wavelength routed optical network", Proc 18th European Conf on Optical Communications, Berlin, Germany, (1992), p 609

[3] J E Midwinter, "Photonics in Switching: The next 25 yearsof optical communications?", IEE Proceedings-J, vol 139, (1992), p 1

[4] H S Hinton, "Overview of freespace photonic switching", Proc Int Top Meeting Photonic Switching, Kobe, Japan, (1990), paper 12 D1

[5] M N Islam, C J Chen, C E Soccolich, "Ultrafast all-optical fiber soliton logic gates", Proc Top Meeting Photonic Switching, Salt Lake City,USA, (1991), paper WD2

[6] D J Blumenthal, K Y Chen, J Ma, R J Feuerstein, J R Sauer, "Demonstration of a deflection routing 2 × 2 photonic switch for computer interconnects", IEEE Phot Techn Lett, vol 4, (1992), p 169

[7] P Prucnal, P Perrier, "Self routing photonic switching with optically processed control", Optical Engineering, vol 29, (1990), p 170

[8] T. M. Martinson, "Synchronization of passive access nodes in very high speed optical packet networks", Proceedings of ECOC '90, Amsterdam, The Netherlands, p 473, (1990)

[9] J M Gabriagues, "Exploitation of the wavelength domain for photonic switching in the IBCN", Proc 17th European Conf on Optical Communications, Paris, France (1992), p 59

[10] M Eiselt et al,"One million pulse circulations in a fiber ring using a slalom for pulse shaping and noise reduction", Proc 18th European Conf on Optical Communications, Berlin, Germany, (1992), p 875

[11] S. Johansson, M. Lindblom, P. Granestrand, B. Lagerström, L. Thylén, "Optical Cross-Connect Systems in Broadband Networks: System Concept and Demonstrators Description."; G. R. Hill, et. al. "A Transport Network Layer Based on Optical Network Elements.", to be published in J Lightwave Techn.

[12] T Shiragaki et al, "An experimental optical digital crossconnect system using $LiNbO_3$ photonic switch matrices and optical amplifiers", Proc 18th European Conf on Optical Communications, Berlin, Germany, (1992), p 605

[13] M. Gustavsson, L. Thylén, A. Djupsjöbacka, *Electron.Lett.*, 25: 1375, (1989)

[14] L. Thylén, T Palm, to be published

Transition from Optical Interconnections to Optical Computing

Richard C. Williamson
Massachusetts Institute of Technology
Lincoln Laboratory
Room C317, 244 Wood St.
P.O. Box 73
Lexington MA 02173-6479

Optical interconnections will be a foot in the door to the next generation of computer systems and provide an evolutionary path toward the use of optics in finer scales.

Optical Computing Posters: 1

OTuA 6:30pm-8:00pm
Grand Ballroom Center

CONCEPT AND IMPLEMENTATION OF THE FRACTIONAL FOURIER TRANSFORM

Adolf W. Lohmann
Universität Erlangen-Nürnberg
Angewandte Optik, Staudtstr.7/B2
8520 Erlangen, Germany
Tel. 49-9131-858371; Fax 49-9131-13508

We assign the degree P=1 to the ordinary Fourier transform. The fractional Fourier transform, for example with degree P=1/2, performs an ordinary Fourier transform, if applied twice in a row. Mendlovic and Özaktas did introduce the fractional Fourier transform into optics [1] based on the fact that a piece of GRIN fiber of proper length will perform a Fourier transform. Cutting that piece of GRIN fiber into shorter pieces corresponds to splitting the ordinary Fourier transform into fractional transforms. We approach the subject of fractional Fourier transforms in two other ways: first by pointing out the algorithmic isomorphism between rotation of the Wigner distribution function and fractional Fourier transforming. Second, we propose two optical setups that are able to perform a fractional Fourier transform.

What does this transform have to do with optics ?

The (ordinary) Fourier transform is of such central significance to physical optics and to optical information processing that everything which is somehow related to Fourier mathematics is likely to be important as well in the realm of optics.

What is a fractional Fourier transform ?

Every image u(x,y) or signal u(t) can be described indirectly and uniquely by a WIGNER distribution function [2,3]. The Wigner distribution function WDF undergoes certain changes if something happens to the signal (from now on called u(x)). For example, propagation in free space means a horizontal shearing of the WDF and passage through a lens corresponds to a vertical shearing of the WDF. A Fraunhofer diffraction (i.e. an ordinary Fourier transform) lets the WDF rotate by 90°. Hence it is plausible to define a fractional Fourier transform as what happens to the signal u(x) while the WDF is rotated by an angle of $\phi = P\pi/2$. The P is the " fractional degree ". Notice, two consecutive rotations obey $\phi_1+\phi_2=\phi_{TOTAL}$ and $\phi_1+\phi_2=\phi_2+\phi_1$. Hence, our definition is inherently additive and commutative.

How can one implement experimentally the fractional Fourier transform ?

With three shearing operations in cascade one can perform a rotation of an image (fig.1). Since we know how to shear a Wigner distribution function WDF, we also know how to rotate the WDF. The three proper shearing operations upon the Wigner distribution function correspond to " free space - lens - free space ", as shown in fig.2 . As an alternative, the three steps may be "lens - free space - lens" (fig.3). The first setup (fig.2) demands

$$Z = f_1 \tan(\phi/2) \quad ; \quad f = f_1/\sin(\phi) \qquad (1)$$

And for the second setup one has to chose :

$$f = f_1 \tan(\phi/2) \quad ; \quad Z = f_1/\sin(\phi) \qquad (2)$$

Notice, the special case of $\phi = 90°$, i.e. the ordinary Fourier transform (P=1) , emerges from eq.1 and 2. And the two setups are now the standard Fourier transform units.

It is a great pleasure to acknowledge many stimulating discussions with David Mendlovic and Haldun M. Özaktas. They introduced me to the subject of fractional Fourier transforms.

References

(1) H.M.Özaktas, D.Mendlovic, "Fourier transforms of fractional order and their optical implementation", **subm. to Opt.Lett.**
(2) M.J.Bastiaans, "The Wigner distribution function applied to optical signals and systems", **Opt.Comm.25(1978)26-30.** "Wigner distribution function and its application to first order optics", **J.Opt.Soc.Am.69(1979) 1710-1716.**
(3) H.O.Bartelt, K.-H.Brenner, A.W.Lohmann, "The Wigner distribution function and its optical production", **Opt.Comm.32(1980) 32-38.**

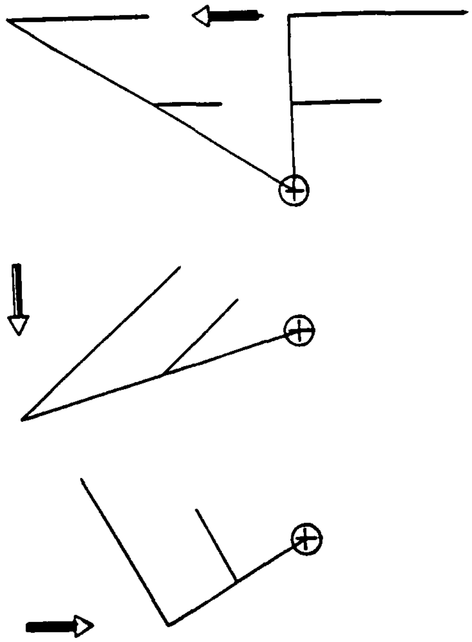


Fig.1: Image rotation, generated by three shearing processes: to the left; down; to the right.

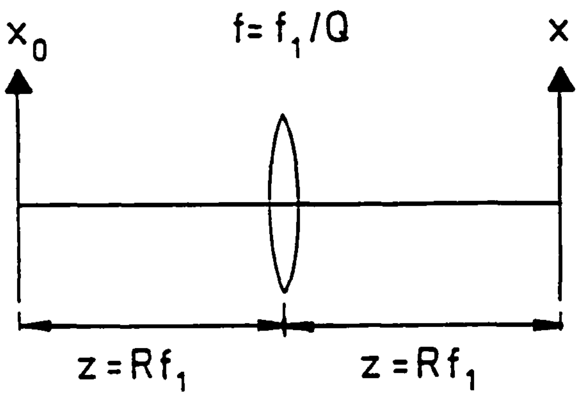


Fig.2: Setup (type I) for performing a fractional Fourier transform. Parameters R and Q determine the degree P and the angle $\phi = P\pi/2$.

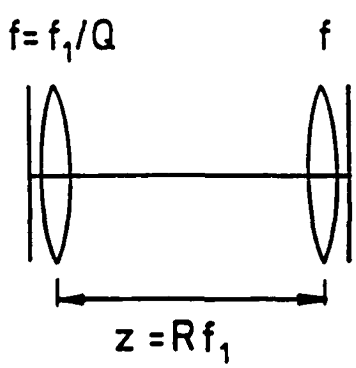


Fig.3: Setup (type II) for performing a fractional Fourier transform.

ITERATIVE DESIGN OF COMPUTER GENERATED FRESNEL HOLOGRAMS FOR FREE-SPACE OPTICAL INTERCONNECTIONS

Bernard C. KRESS, Sing H. LEE
Department of ECE, University of California at San Diego, La Jolla, CA 92093-0407
Tel:(619)534 2325 Fax:(619)534 1225

I. Introduction

We present an iterative algorithm to design computer generated holograms (CGHs) for optical interconnections. The performances are compared to those of similar elements generated by several non-iterative methods. Parameters of comparison include those of input/output characteristics (semiconductor laser beam profiles, uniformity of reconstruction, bit error rate) and those of fabrication (minimum feature sizes, calculation time and costs). To show how well the iterative algorithm works, we apply it to design CGHs to be used in free-space optically interconnected multiprocessor parallel computing systems based on twin butterfly architectures.

II. Iterative algorithm

The iterative design method used here is an error-reduction algorithm known as the Gerchberg Saxton (G-S) algorithm [1]. The algorithm generates a phase only CGH given a set of input and ouput amplitude distributions (laser beam profile and desired interconnection pattern). The algorithm begins by generating a random phase mask. This initial phase function is multiplied by the amplitude distribution of the incomming beam to yield a complex wavefront in the CGH plane. A Fresnel transform propagates this wavefront to the processing element (PE) plane, where the resulting amplitude is replaced by the interconnection pattern, while the phase distribution is kept. An inverse Fresnel transform propagates then this new wavefront back to the CGH plane. Again, the phase distribution is kept and the amplitude function is replaced by the laser beam amplitude profile, and another iteration begins.

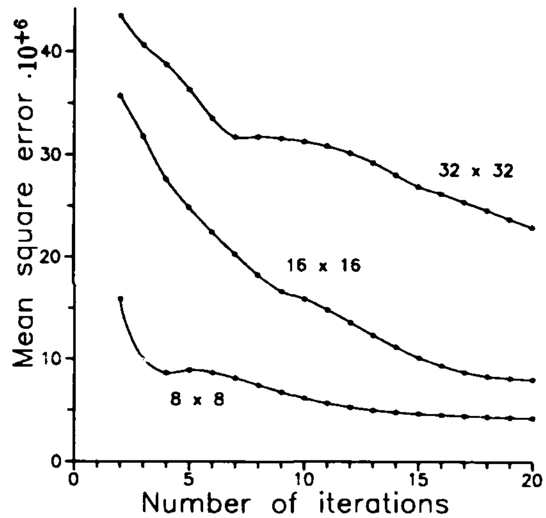


Fig.1: *Convergence of the algorithm for different array sizes.*

The root mean square error is a good parameter to evaluate the convergence rate of the algorithm. In Fig.1 we calculated 3 CGHs for increasing fan-out array dimensions. The convergence rate was found to be closely related to the complexity of the arrays.

III. Non-iterative design methods.

For comparison, we selected the three following non-iterative methods. Our iterative method will be referred to as the 4^{th} method.

1o **FZP array.** A CGH composed of an array of FZPs is described here. They are easy to calculate and fabricated as phase only CGHs [2]. For N facets in the array, the wavefront to encode is:

$$e^{i\Phi_s(x,y)} = \sum_{n=0}^{N} Rect\left(\frac{x-x_n}{X}\right) Rect\left(\frac{y-y_n}{Y}\right) e^{-i\Phi_n(x,y)}\ ,\ \Phi_n(x,y) = \frac{\pi}{\lambda f}[(x-x_n)^2+(y-y_n)^2],$$

where x_n and y_n are the offsets in both directions and X and Y the facet dimensions.

2o Spatially multiplexed FZPs. The different FZPs are added over the same CGH space [3]. The resulting amplitude is no longer unity, and the complex wavefront to encode becomes:

$$\mathcal{A}_m(x,y)\cdot e^{i\Phi_m(x,y)} = \sum_{n=1}^{N} e^{-i\Phi_n(x,y)}$$

The special case where the amplitude is set to unity is case 2a and requires phase only encoding, whereas the general case is 2b and requires complex encoding.

3o Spatial modulation of Fourier CGH by a FZP. This method requires two steps [4]: First, a Fourier CGH is calculated to perform the required fan-out in the far-field. Second, this CGH function is modulated by a FZP.

$$\mathcal{A}_f(x,y)\cdot e^{i\Phi_f(x,y)}\cdot e^{\frac{-i\pi}{\lambda f}(x^2+y^2)}$$

where $\mathcal{A}_f$ and Φ_f are the Fourier CGH functions, and $e^{\frac{-i\pi}{\lambda f}(x^2+y^2)}$ the FZP.

IV. Performances simulation of the different designs.

A twin butterfly interconnection architecture has been chosen to simulate and compare the performances of the different design techniques. To implement this architecture, we need to make an array of CGHs, each providing four irregular interconnections (e.g. a fan-out of four). Fig.2 shows the diffraction efficiencies for CGHs of different number of phase levels for the 4 designs. The simulations assume perfect etch depths, alignments, perfectly square cells and uniform illumination. The signal-to-noise ratio ($\mathcal{SNR}$) shown in Fig.3 is an important characteristic of the CGH since it is directly related to the bit error rate. Fig.4 shows the intensity profiles of the reconstructions from the different one to four fan-out CGHs design methods. The first case has good diffraction efficiency but poor $\mathcal{SNR}$, and since each FZP has only a fraction of the CGH area, the spots become larger, reducing the $\mathcal{SNR}$ (Fig.4.a). Spatially multiplexing FZPs can yield to very good $\mathcal{SNR}$ when the entire complex function is encoded in the CGH, (case 2b), but the resulting diffraction efficiency drops dramatically. The diffraction efficiency can be increased at the expenses of $\mathcal{SNR}$ by encoding the phase only (case 2a). The diffraction limited spots are small because the different FZPs occupy the entire CGH space (Fig.4.b). Cases 4 gives the best diffraction efficiency and $\mathcal{SNR}$.

The imaging qualities of the CGH in case 3 are deteriorated by aberrations due to the focussing of all the spots by a single FZP (Fig.4c). This can be avoided in case 2b where each FZP can be optimally designed to focus each single spot. The sharp uniform spots in case 4 predict good imaging qualities for this design method.

We have also examined the effects of illumination uniformity on the intensity uniformity of the fan-out spots. In Fig.5 we simulated the fan-out reconstruction for different input beam waists of an assymetric TEM_{00} configuration, typically a laser diode beam (the beam waist is taken along the major axis). Uniformity in case 1 decreases very rapidly with the edge facets diffracting very low intensity compared to the center facets. Case 2.b and 3 have the fan-out information spread all over the CGH area: this yields higher uniformity than in previous case, but still less than 50% for a ratio of 1/2, (e.g. beam intensity profile truncated at $1/e^2$). Case 4 has been designed for this particular beam waist: the uniformity remains relatively high ($\simeq$ 87%) for the considered beam waist before dropping as the other cases.

Finally, the CPU time required for the different CGH designs (Fig.6) is considered. Case 4 requires long CPU time (depending on the number of desired iterations). Design in case 3 only performs a single FFT and a FZP modulation, whereas in case 2b all the FZPs have to be calculated (and optimized) over the entire CGH space. The 1^{st} case is the fastest, since each FZPs is calculated over a single facet.

V. Fabrication considerations.

An important parameter in the fabrication process is the minimum feature size. Fig.7 shows the minimum feature sizes for different CGHs (for the same interconnection architecture than in IV). These values have been determined by evaluating the derivative of the phase function. In 3^{rd} and 4^{th} cases the frequencies have been scanned through the calculated CGH. The fabrication costs of such a CGH are directly related to these feature sizes, since the e-beam spot size can be enlarged to plot large features. If the enlargement is x times in both directions, the e-beam writing time is decreased by a factor of approximately x^2 (x takes only some discrete values). The e-beam spot size was chosen to be at least half of the minimum feature size. Case 1 yield relatively large feature sizes when compared to cases 2 and 4. The very small feature sizes in case 2b result from the complex encoding.

VI. Experimental results.

Our iterative algorithm as well as the three non-iterative design methods have been applied to the design of CGH arrays for twin butterfly interconnection architectures. The 4 different CGH arrays have been realised with electron-beam technology by direct write of 16 phase levels into positive e-beam resist. Figure 8 shows for each design 4 CGHs out of the 8x8 twin butterfly CGH array for the first stage of the twin butterfly network.

VII. Conclusion.

We have investigated an iterative method to calculate focussing fan-out CGHs and compared them to non-iterative design methods. It was found that this method provides the best compromise between performance and fabrication considerations.

References.

[1] **R.W.Gerchberg and W.O.Saxton.** "A practical algorithm for the determination of phase from image and diffraction plane pictures". *Optik*, **35**, *No.2, 237-246 (1972).*

[2] **J.Jahns et al.** "Integrated free-space optical permutation network" *Technical digest, Topical meeting on optical computing*, **6**, *29-31 (1991).*

[3] **H. Zarschizky et al.** "Clock distribution using a synthetic HOE with multiple fan-out" *SPIE* **1281**, *103 (1990).*

[4] **Y. Hatakoshi et al.** "Phase-modulated grating lenses for two dimensional optical branching" *Jpn. J. Appl. Phys.* **31**, *1606-1610 (1990).*

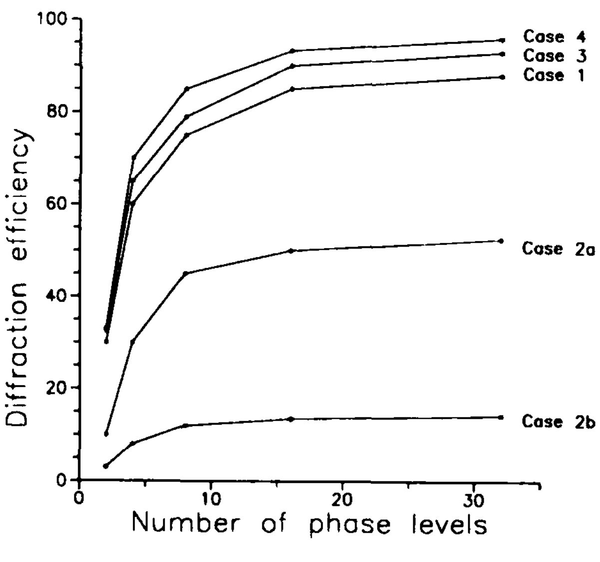


Fig.2: *CGH diffraction efficiencies versus number of phase levels.*

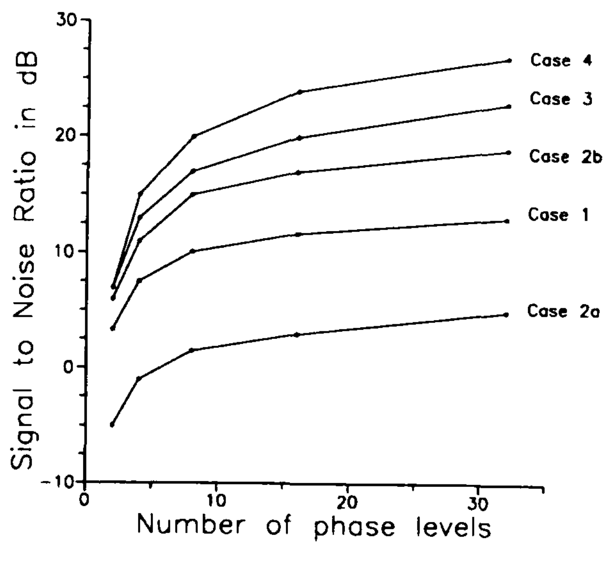


Fig.3: *Signal to Noise Ratio versus number of phase levels.*

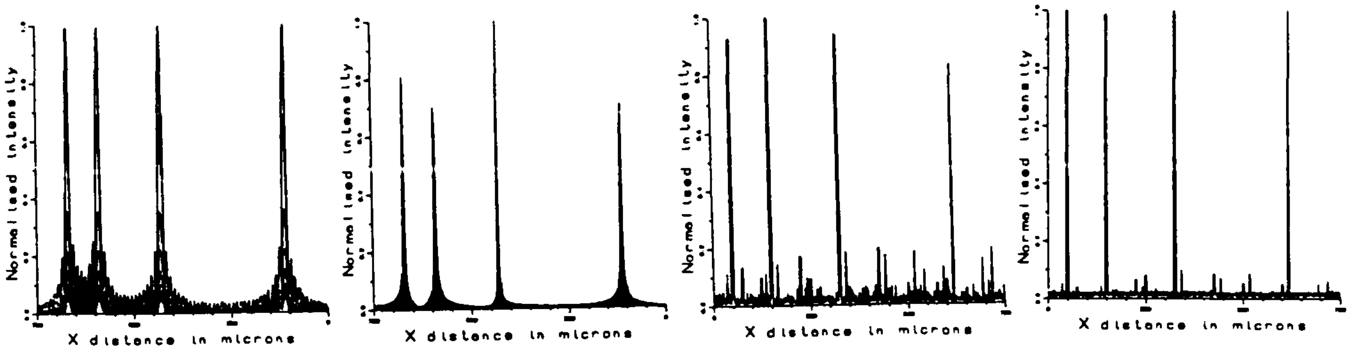


Fig.4: *Reconstructions of an irregular 1 to 4 fan-out interconnection (11010001) from the different CGH design methods. (a) case 1, (b) case 2b, (c) case 3, (d) case 4.*

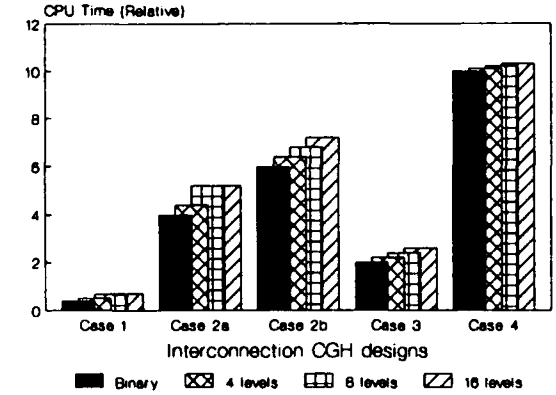


Fig.6: *CPU time required for the different CGH designs.*

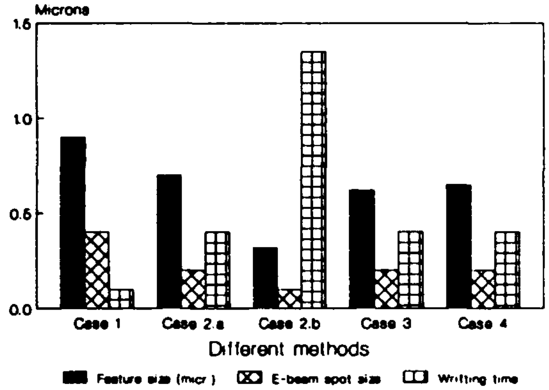


Fig.7: *Minimum feature sizes, e-beam spot size and writing time for the different CGHs.*

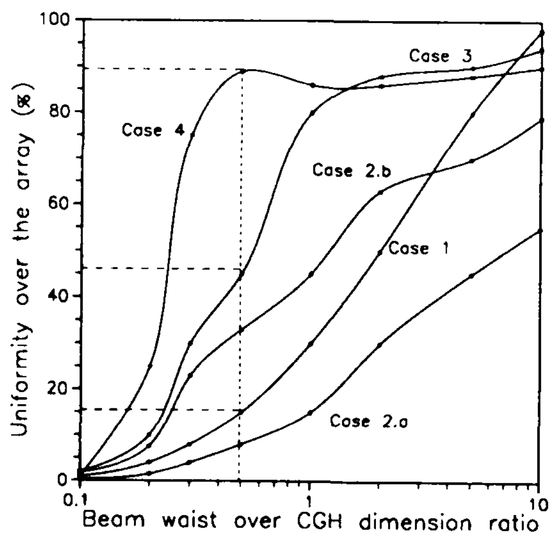


Fig.5: *Uniformity of the reconstruction for different input beam waists (TEM_{00} mode).*

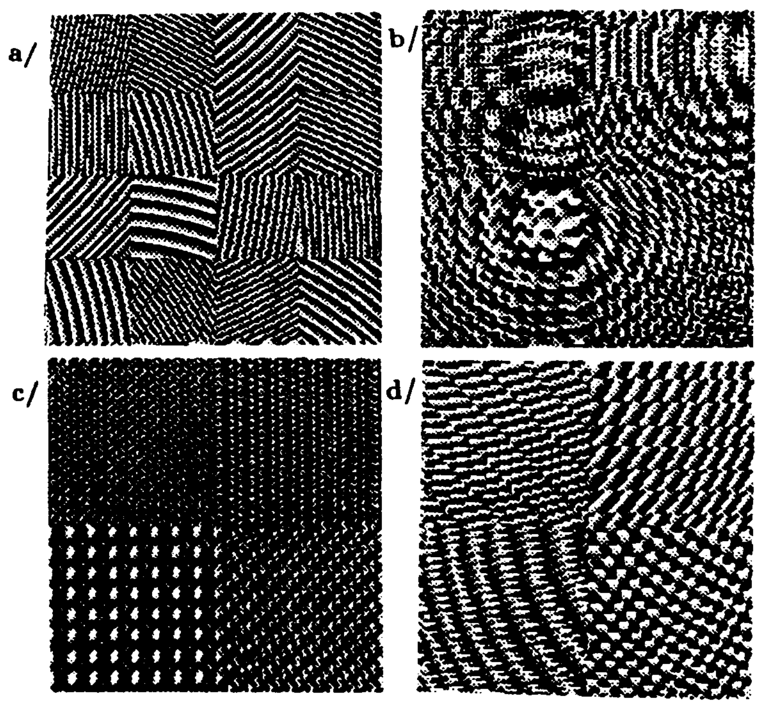


Fig.8: *Optimized twin butterfly interconnection module for the first layer (4 CGHs out of the 8x8 array) for the 4 different CGH design techniques. (a) case 1, (b) case 2a, (c) case 3, (d) case 4.*

Crosstalkless Incoherent Optical/Electronic Hybrid Associative Memory System

Masaki Taniguchi, Katsunori Matsuoka†, and Yoshiki Ichioka

Department of Applied Physics, Faculty of Engineering, Osaka University
2-1 Yamadaoka, Suita 565, Japan
†Government Industrial Research Institute, Osaka
1-8-31 Midorigaoka, Ikeda 563, Japan

Introduction

An associative memory is a fault-tolerant content addressable memory system. It can be applied to restoration or substitution of patterns. The concept of associative memory has been developed in two disparate fields, namely, neural networks and holography. Associative memories[1,2] using optical correlators can directly process two-dimensional information. They have lately attracted attention in the field of optical information processing.

Recently, we have proposed and developed a new type of optical associative memory system[3,4] using incoherent optical correlators. The proposed associative memory system consists of incoherent optical correlators and video devices. The system has the following features:

(1) By using incoherent correlators, the system is easy to compose and to handle as compared with using coherent ones. In addition, it is easy to produce correlation filters.
(2) Either a binary or gray-level pattern can be processed and associated.
(3) Both auto-association and hetero-association can be carried out.

In the proposed system, multiple-object discriminant correlation filters are used to classify an input pattern. The filters are calculated from a set of prepared training patterns. It is able to specify the correlation values between filters and training patterns. Synthetic discriminant function (SDF)[5] filters are typical multiple-object discriminant filters, and they have been used in the proposed system[3,4]. However, there is a problem when SDF is recorded on a medium like a photographic film. SDF has values with a wide range and very gentle gradation. Therefore, if the gradation of the medium is not sufficiently gentle, SDF cannot be accurately recorded. This causes error of correlation response of the filter. As a result, the error causes the crosstalk in the associative memory.

We calculate multiple-object discriminant filters by extended simulated annealing algorithm[6]. They show the expected correlation responses. By using the filters for our associative memory system, the crosstalkless association results are obtained.

Optical/electronic hybrid associative memory system

Figure 1 is a schematic diagram of proposed optical/electronic hybrid associative memory system.

The input pattern on the monitor TV is imaged on CCD camera (CCD1) and displayed on liquid crystal television (LCTV1). The pattern is correlated with the array of discriminant filters, and then the correlation result is detected by CCD camera (CCD2). The filter array consists of discriminant filters corresponding to the memorized classes. Each filter is designed so as to classify the input pattern to one of the classes according to the correlation peak value. The correlation image obtained by CCD2 is displayed on LCTV2. A mask in front of LCTV2 picks out the correlation peaks corresponding to the individual memorized classes. Since LCTV2 has a nonlinear I/O response, the correlation peaks obtained CCD2 are nonlinearly transformed into brightness.

Bright spots displayed on LCTV2 through the mask project the array of recall filters. The array consists of memorized patterns to be recalled corresponding to individual memorized classes. Individual recalled patterns are piled up and superposed on the input pattern on CCD1.

The recalled pattern consists of patterns corresponding to more than two classes, because correlation peaks for some classes have high values. To recall a crosstalkless output pattern, it is necessary to leave only one correlation peak on CCD2, and to suppress other unnecessary correlation peaks. In the system, these operations can be achieved by the following processing.

The intermediate processed image on CCD1 during iteration is re-displayed on LCTV1, and re-correlated with the array of discriminant filters. Owing to tthe nonlinear characteristic of LCTV2, the highest correlation peak is amplified more than other correlation peaks. The CCD camera (CCD3) detects the total power of bright points on LCTV2. The signal of total power controls the diaphragm of auto-iris-lens (AIL1) so as to keep the total power to be constant. By this processing, the highest correlation peak is left after convergence. Thus, a crosstalkless recalled pattern corresponding to the highest correlation peaks is obtained on CCD1.

Array of discriminant filters

To attain crosstalkless association by the nonlinear feedback method mentioned above, the discriminant filters must satisfy two following conditions:

(1) For a pattern which belongs to one of memorized classes, the filter corresponding to the class outputs higher correlation peak value than filters of other classes.

(2) For each memorized pattern to be recalled, the filter corresponding to the pattern outputs an specified correlation peak value, and other filters output zero correlation peak value.

As the filter function which satisfies above conditions, SDF filter is typical. We have used SDF for the array of discriminant filters. However, SDF has values with a wide range and very gentle gradation. Therefore, it is difficult to record SDF on a recording medium accurately. This inaccuracy causes error of the correlation response of the filter, and crosstalk is occured on a recalled pattern.

To get the crosstalkless output pattern, we use simulated annealing algorithm to generate a new filter function with few steps of gray-level. The filter function provide a better correlation response than SDF. To calculate the filter functions with a few steps of gray-level makes it easy to control the actual filter function on photographic film.

Experiments

Figure 2 shows experimental results by the proposed associative memory system. Training patterns are shown in Fig.2(a). In the experiments, we memorize four classes in the system. The memorized classes are "Circle," "Rectangle," "Star," and "Triangle," respectively. The memorized patterns to be recalled are the first letters of the words. Figure 2(b) shows input patterns, which are defective or corrupted by noise. The results of association are shown in Fig. 2(c). It is shown that the memorized pattern of the correct class is recalled for each input pattern The association time is about 0.8 second. That is limited by the response time of diaphragm of auto-iris-lens.

In the proposed system a recalled pattern corresponds to the highest peak that appeared on the correlation plane of the discriminant filters and an input pattern. A winner-take-all on the correlation plane can avoid the iteration. However, the proposed system functions as the winner-take-all with no thresholding and no comparing. An output pattern is obtained stably without crosstalk independent of an input pattern.

The association time is restricted by the response time of diaphragm of auto-iris-lens. If we can use the auto-iris-lens driving with more high response time, the system can work at TV video rate.

References

1. E. G. Paek and D. Psaltis, "Optical associative memory using Fourier transform holograms," Opt. Eng. 26, 428 (1987).
2. B. H. Soffer, G. J. Dunning, Y. Owechko and E. Marom, "Associative holographic memory with feedback using phase-conjugate mirrors," Opt. Lett. 11, 118 (1986).
3. M. Taniguchi, K. Matsuoka and Y. Ichioka, "Hybrid associative memory using an incoherent correlation system," Jpn. J. Appl. Phys. 29, L1843 (1990).
4. M. Taniguchi, K. Matsuoka and Y. Ichioka, "Incoherent optical associative memory by using synthetic discriminant function filters," Appl. Opt. 31, 3295 (1992).
5. D. Casasent, "Unified synthetic discriminant function computational formulation," Appl. Opt. 23, 1620 (1984).
6. M. Taniguchi, K. Matsuoka and Y. Ichioka, "Multiple-object discriminant filter generated by simulated annealing," Jpn. J. Appl. Phys. 31, 2133 (1992).

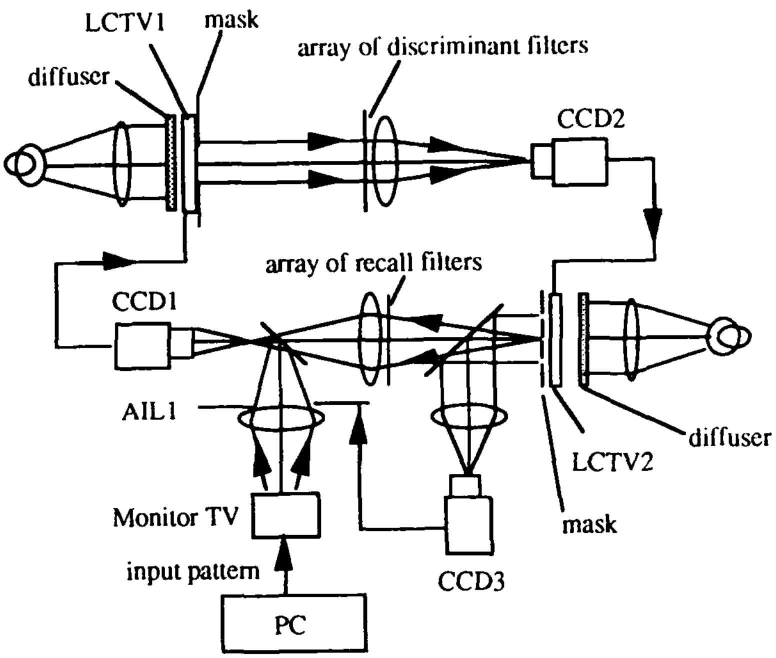
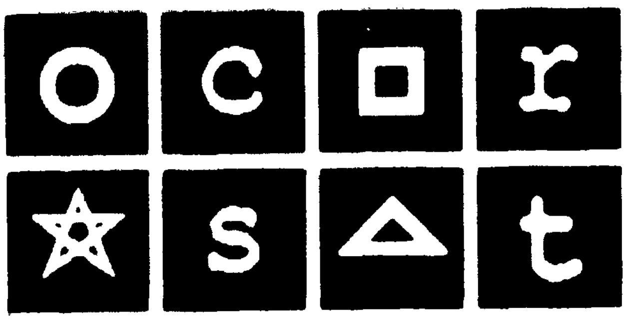
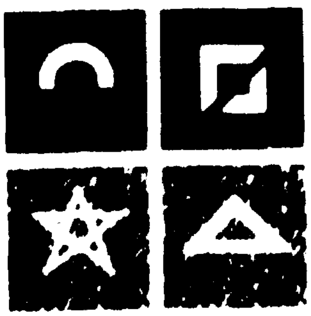


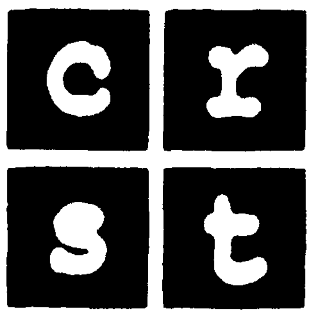
Fig.1 Hybrid associative memory system



(a)



(b)



(c)

Fig.2 Association results (a) training patterns (b) input patterns (c) recalled patterns

Reflective block optics for optical computing systems

Daisuke Miyazaki, Jun Tanida, and Yoshiki Ichioka
Department of Applied Physics, Faculty of Engineering
Osaka University
2-1 Yamadaoka, Suita 565, JAPAN

1. Introduction

To construct an optical computing system, it is important to develop feasible technique to assemble a large amount of optical components into a complex system. Recently, several ideas for three-dimensional (3-D) integrated optics have been proposed to provide high stability, high reliability, ease of alignment, and compactness for optical systems.[1-4]

In most of the 3-D integrated optics, micro lenses are fabricated by the lithography or ion exchange technique, and their diameters are from 10^{-3} to a few millimeters. However, an optical system decreases its space-bandwidth product as decreasing the physical size.[5] In addition, optical interconnections are advantageous over electronic interconnections when the distance of interconnects is long in the respect of power and speed consideration.[6] Thus, it is considerable to develop a technique for optical integration in the range larger than a few millimeters.

In this paper, we propose reflective block optics as a new concept for assembly of reliable optical systems for parallel optical computing. A discrete correlator is designed and several basic experiments are executed to verify the concept.

2. Concept of the reflective block optics

The concept of the reflective block optics is similar to that of the solid optics in which air space in a conventional optical system is fulfilled with solid medium such as glass to obtain a rigid optical system.[7] Larger optical systems can be fabricated by the solid optics than the 3-D integrated optics mentioned above. However, the solid optics has a disadvantage in difficulty to achieve high refracting power, because the difference between the refractive indexes of different kinds of solid media is not large.

In the reflective block optics, reflective lenses provide high power usage, compactness, and ease of fabrication. Figure 1 shows an example of 4-f system using the reflective block optics. To obtain the function of the lenses, concave mirrors made of plano-convex lenses coated with reflective layer are contacted on a cube polarizing beam splitter (PBS). The light propagates through the homogeneous solid media only. To construct a more complex system, several kinds of optical blocks are connected in cascade. Linearly polarized light can be propagated without power loss with the PBS and the quarter wave plates.

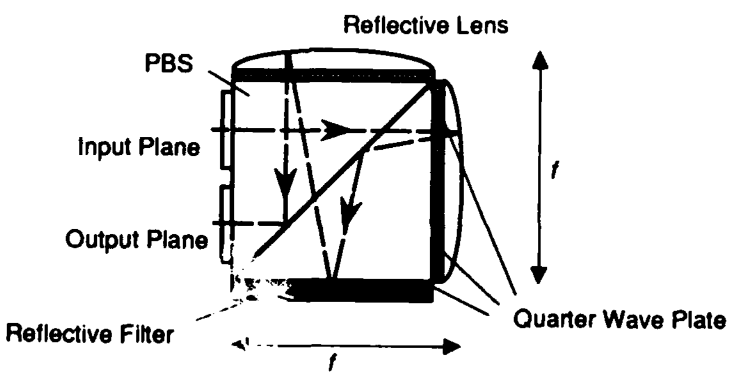
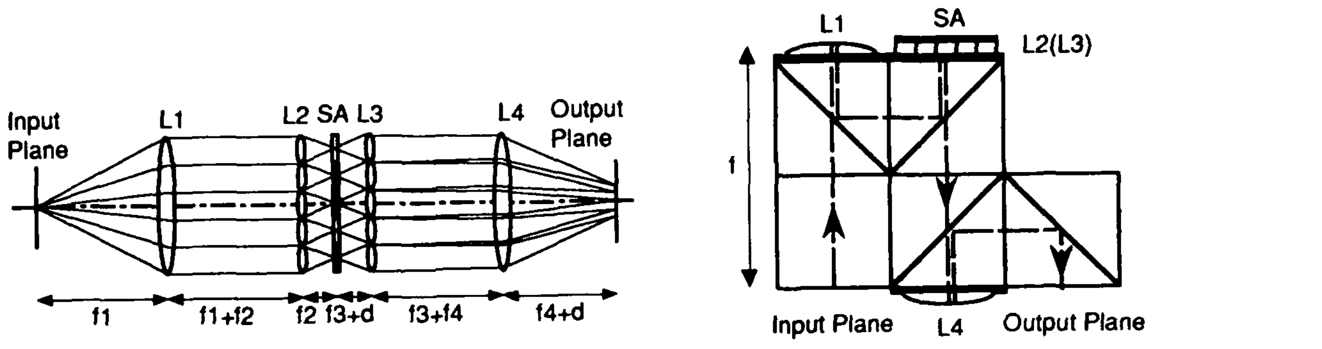
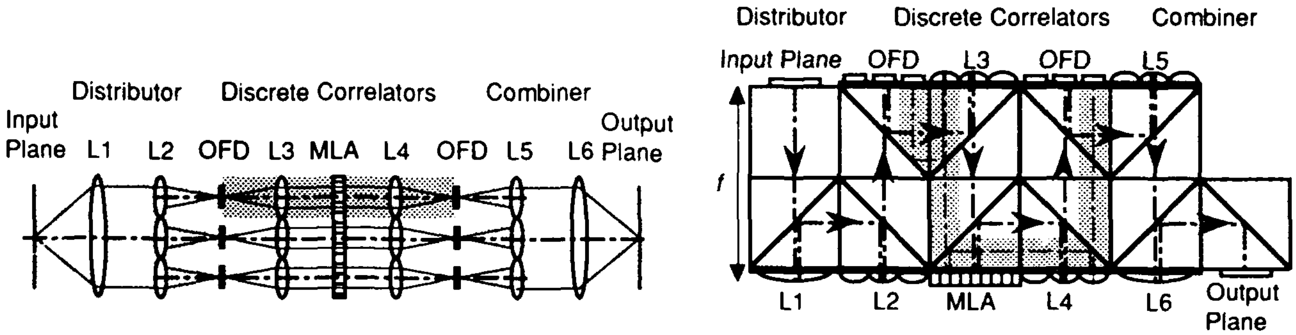


Fig. 1 4-f system by the reflective block optics.



(a) Conventional optical system (b) Assembly usi.g the reflective block optics

Fig. 2 Discrete correlator using multiple imaging with lens arrays. L2, L3: Lens Array, SA: Shatter Array.



(a) Conventional Optical System (b) Assembly using the reflective block optics

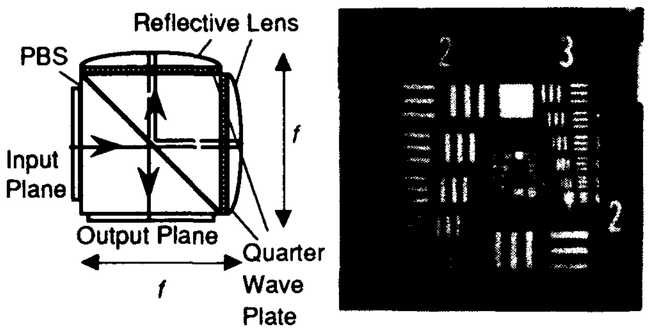
Fig. 3 Optical computing system consisting of parallel discrete correlators for optical array logic and symbolic substitution logic. L1-L6: lens, OFD: optical functional devices, MLA: micro lens array.

3. Applications to digital optical computing systems

Discrete correlation is one of the basic operations in the digital optical computing techniques such as optical array logic [8] and symbolic substitution logic.[9] Figure 2(a) shows an example of the discrete correlator using multiple-imaging with lens arrays.[10] The correlator can be constructed with reflective block optics as shown in Fig. 2(b). For arbitrary logical operation in both of optical array logic and symbolic substitution logic, multiple discrete correlation is required for a single input image. An optical system for the purpose is shown in Fig. 3(a), which consists of a distributor, several discrete correlators, and a combiner. The system executes multiple discrete correlations simultaneously and effectively processes parallel operations based on optical array logic and symbolic substitution logic. Figure 3(b) shows the system constructed with reflective block optics.

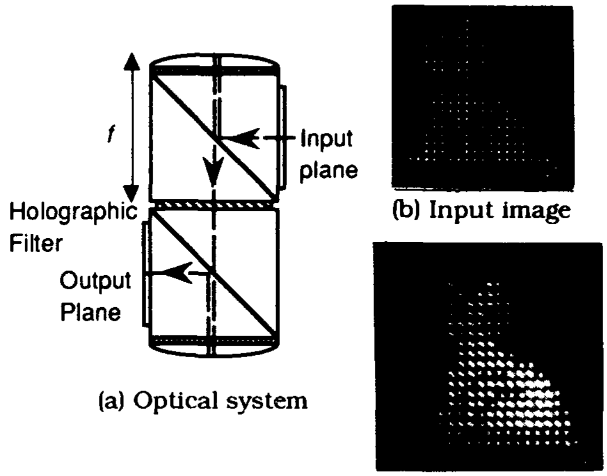
4. Basic experiments

Figures 4 and 5 show experimental results of imaging and discrete correlation with a holographic filter by the reflective block optics. The reflective lenses are made of convex lenses coated by aluminum. The focus length and the diameter are 25.95mm and 20mm, respectively. The size of input image is 5mm x 5mm. On the holographic filter, two gratings oriented to the different directions are recorded by the multiple exposure technique. By the image of the test chart in Fig. 4(b), resolution of the constructed system is estimated as 16 line-pair/mm. Figure 5(c) shows the desired result in which the input image is split into two, and the images are shifted and overlapped.



(a) Optical System (b) Output Image

Fig. 4 Experimental result of imaging by the reflective block optics.



(a) Optical system

(b) Input image

(c) Output image

Fig. 5 Experimental results of holographic discrete correlator by the reflective block optics.

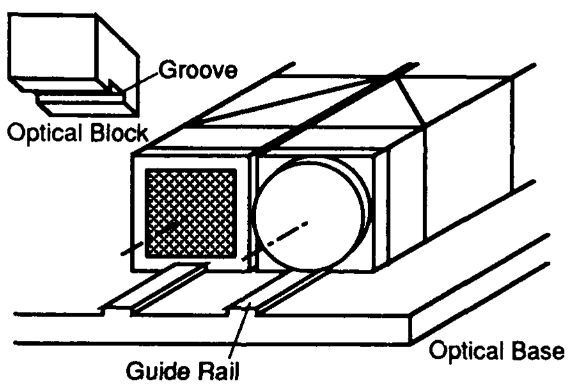
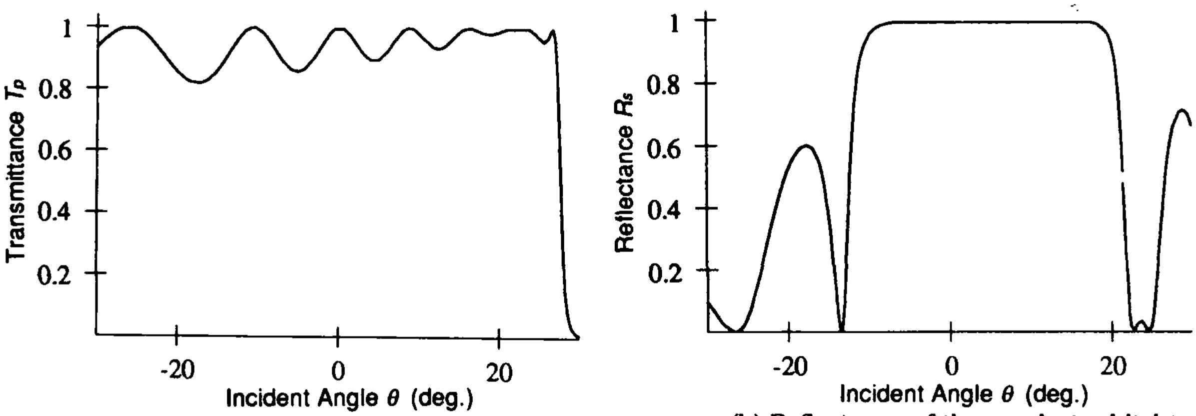


Fig. 6 Alignment system using guide rails and grooves.

5. Discussion

A. Alignment method An easy and precise mechanism is required to align the optical blocks. Figure 6 shows an example of the alignment system. The groove of an optical block is held by a guide rail on an optical base, so that every components are aligned along the guide rail. The length along the optical axis is regulated by the size of optical blocks. However, error of the length is accumulated as the number of cascaded blocks increases. To compensate the error, adjusting blocks are inserted.

B. Restriction of incident angle into polarizing beam splitter Since the reflective block optics is based on PBSs, its allowable incident angle is restricted by the performance of the PBSs.[11] Figure 7 shows the calculated result of the dependence of the transmittance of the p-polarized light and the reflectance of the s-polarized light on the incident angle into a typical PBS with multilayer. From Fig. 7, assuming that the transmittance and reflectance are acceptable at more than 85%, the incidence angle within ±10% is utilized in the optical systems. It is possible to design a PBS offering larger range of incident angle.[12]



(a) Transmittance of p-polarized light   (b) Reflectance of the s-polarized light

Fig. 7 Example of the dependence of the transmittance of p-polarized light and the reflectance of s-polarized light on incident angle into a PBS. The design of multilayer coating is $(L/2\ H\ L/2)^5$ with $n_p=1.619$, $n_H=2.05$, $n_L=1.38$.

*C. **Stray light*** During the experiments of the reflective block optics, ghost images are observed on the output image. To avoid such stray light, several techniques should be applied; for example, cutting the undesired light with a stop, suppressing the internal reflection at the side of blocks with anti-reflecting coat, and matching the refractive indexes at the block connection.

6. Conclusions

A new concept for optical assembly, called the reflective block optics, has been proposed to achieve stability, reliability, compactness, and simplification in alignment of optical computing systems. Using the reflective block optics, optical systems whose space-bandwidth product is larger than 3-D integrated micro optics can be constructed. Imaging and discrete correlation have been demonstrated to verify the concept of the reflective block optics. The resolution is observed as 16 line-pair/mm in the experiment. The incidence angle within ±10% is allowable in the reflective block optics, assuming that the transmittance and reflectance for individual polarizing components of a PBS is more than 85 %.

As a future work, reliability related on fabricating precision and signal crosstalk must be studied as well as construction techniques for more complex optics.

References

[1] K. Iga, M. Oikawa, S. Misawa, J. Banno, and Y. Kokubun, "Staked planar optics: an application of the planar microlens", Appl. Opt., **21**, 3456-3460 (1982).
[2] K.-H. Brenner, "3D-integration of digital optical system," Optical Computing, 1991, Technical Digest Series, **6**, 25-28 (Optical Society of America, Washington, DC, 1991).
[3] J. Jahns and A. Huang, "Planar integration of free-space optical components", Appl. Opt., **28**, 1602-1605 (1989).
[4] K. Hamanaka, "Optical bus interconnection system by using selfoc lenses and planer microlens arrays," Optical Computing, 1991, Technical Digest Series, **6**, 32-35 (Optical Society of America, Washington, DC, 1991).
[5] A. W. Lohmann, "Scaling Laws for lens systems," Appl. Opt., **28**, 4996-4998 (1989).
[6] M. R. Feldman, S. C. Esener, C. C. Guest, and S. H. Lee, "Comparison between optical and electrical interconnects based on power and speed considerations," Appl. Opt., **27**, 1742-1751 (1988).
[7] M. Temmen and C. Hester, "Solid Optical Correlators," Proc. SPIE, Digital Optical Computing II, **1215**, 422-431 (1990).
[8] J. Tanida and Y. Ichioka, "Optical-logic-array processor using shadowgrams. III. Parallel neighborhood operations and an architecture of an optical digital-computing system," J. Opt. Soc. Am., **A2**, 1245-1253 (1985).
[9] K.-H. Brenner, A. Huang, and N. Streibl, "Digital optical computing with symbolic substitution," Appl. Opt., **25**, 3054-3060 (1986).
[10] J. Tanida and Y. Ichioka, "Discrete correlators using multiple imaging for digital optical computing", Opt. Lett., **16**, 599-601 (1991).
[11] R. P. Netterfield, "Practical thin-film polarizing beam-splitters," Optica Acta, **24**, 69-79 (1977).
[12] J. Mouchart, J. Begele and E. Duda, "Modified MacNeille cube polarizer for a wide angle field," Appl. Opt., **28**, 2847-2853 (1989).

CELLULAR TWO-LAYER LOGIC ARRAY AND OPTOELECTRONIC IMPLEMENTATION

Liren Liu, Zibei Zhang, Xuejun Zhang and Jun Zheng

Shanghai Institute of Optics and Fine Mechanics, Academia Sinica, P.O.Box 800-211, Shanghai 201800, P.R.China
Fax:0086-21-9528885 Tel:9528812

INTRODUCTION

One of the strong candidates of the parallel architectures for optical computing is cellular array[1]. Recently, there were some reports on the cellular logic array architectures for binary image processing and their optical implementations[2,3]. As a development of our previous suggestion for the implementation of CLIP architecture[4], in this paper, we propose a new and unified architecture of cellular-morphologic two-layer logic array for binary image processing. The array not only has the image processing functions more than that available by the conventional cellular and morphologic operations but also has the ability to transform images. Moreover, by using the concept of threshold-decomposition, grey-level image processing is also possible. A compact optoelectronic system has been developed for executing the proposed architecture. Theoretical and experimental studies has indicated its power and versatility in image processing.

ARCHITECTURE CONSIDERATION

Fig.1 shows schematically the suggested architecture, which consists of a neighborhood connectivity changeable boolean voting logic layer and a two-channel pattern logic layer. In the processing layer of two-channel pattern logic, the input image A_o and the intermediate morphologic image A_n results in an output image A_1 and an intermediate connection pattern A_2. In the layer of neighborhood boolean voting logic, A_2 fans out into interconnection with N neighboring cells according the required connectivity of tessellation, and then the result is thresholded by a possible level from 1 to N. A voting logic determined morphologic image A_n is thus generated. The procedure is repeated iteratively.

The characteristics of the algorithm are: (1) The voting logic extends the basic morphologic operations into more general ones. (2) The variable neighborhood connections leads to flexibility in morphologic processing. Moreover, image transformation can be realized by changing the connection structures during operations[5]. For example, symbolic substitution and Hough transform can be easily performed. (3) Pattern logic provides the possibility for logic comparison among the image cells and their neighbors. For a determined structure of neighborhood connection and a special threshold of voting logic, there are total 512 functions for image processing. A detailed study shows that only 52 functions are distinct, which are related the immediate connection of neighbors or the margin connection of neighbors. Some of them are conventional mathematical morphologic operations such as erosion, dilation, closing, and opening. (4) Grey-level images can be processed in terms of the threshold-decomposition technique as shown in Fig.2. The input image is thresholded into multiple channels of binary cellular arrays, adding the results from the arrays leads to a processed grey-level image.

OPTOELECTRONIC IMPLEMENTATION

The construction of the optoelectronic 64×64 array system is shown in Fig.3. A quadruple-imaging unit with fourfold-rail spatial coding of input image and polarization customizing of lenses is suggested to carry out two-channel pattern logic operations. By controlling the polarization directions of 4 polarizers, two independent logic operations can be achieved[6,7]. A defocused unit of coded aperture is used for neighborhood interconnection[8,9]. Thresholding is performed by electronics[10]. Fig.4 demonstrates examples of experimental results of erosion and edge detection.

The authors acknowledge the supports by the Chinese Academy of Sciences, the Office of High-Tech, and the National Natural Foundation of China.

References:

1. K.Preston,Jr., M.J.B.Duff, S.Levialdi, P.E.Norgren, and J.Toriwaki, "Basis of Cellular Logic with Some Applications in Medical Image Processing," Proc.IEEE 67, 826-856(1979).
2. K.S.Huang, B.K.Jenkins, and A.A.Sawchuk, "Binary Image Algebra and Optical Cellular Logic Processor Design," Compt.Vision Graphics Image Process. 45, 295-345(1989).
3. A.C.Walker et al, "Design and construction of a Programmable Optical 16x16 Array Processor," Optical Computing(1991 Technical Digest Series Vol.6), 199(1991).
4. L.Liu, X.Liu, and B.Cui, "Optical Programmable Cellular Logic Array for Image Processing," Appl.Opt. 30, 943-949(1991).
5. L.Zhang and L.Liu, "Morphological Operations for Symbolic Substitution and Hough Transform," Opt.Commun. 76, 194-198(1990).
6. L.Liu and X.Liu, "Cascadable Binary Logic Processor Using Multiple-Imaging," Opt.Commun. 82, 446-452(1991).
7. L.Liu, Z.Zhang, and Y.Zheng, "Programmable Optical Parallel Binary Double-Out Logic Processor," Acta Optica Sinica (in press).
8. K.S.O'Neill and W.T.Rhodes, "Morphological Transformations by Hybrid Optical-Electronic Methods," SPIE Vol.638 41-44(1986).
9. X.Zhang, Z.Zhang, and L.Liu, "Mathematical Morphology by an Optical Convolver with Coded Aperture," (submitted).
10. L.Liu, "Optoelectronic Implementation of Mathematical Morphology," Opt.Lett. 14, 482-484(1989).

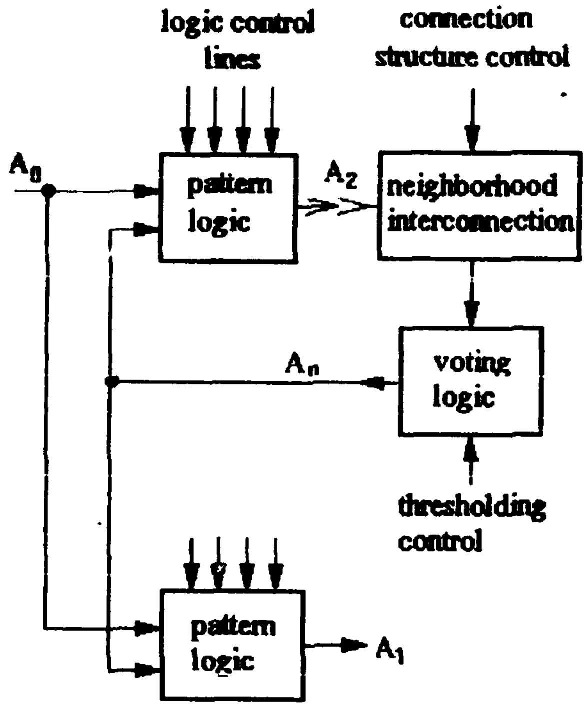


Fig.1. Architecture of cellular-morphologic two-layer logic array for binary image processing

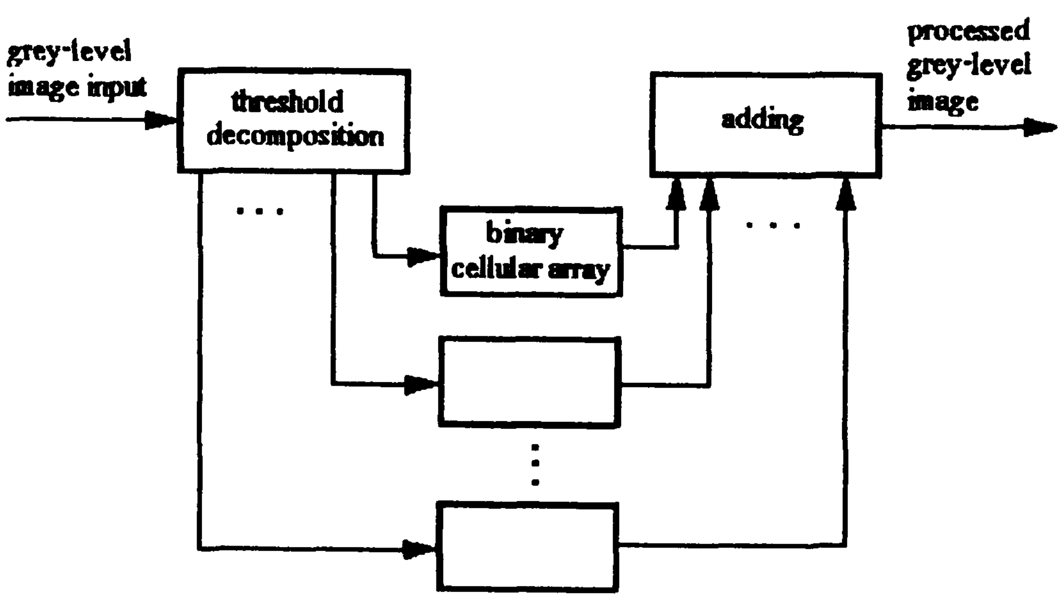


Fig.2. Grey-level image processing by threshold decomposition

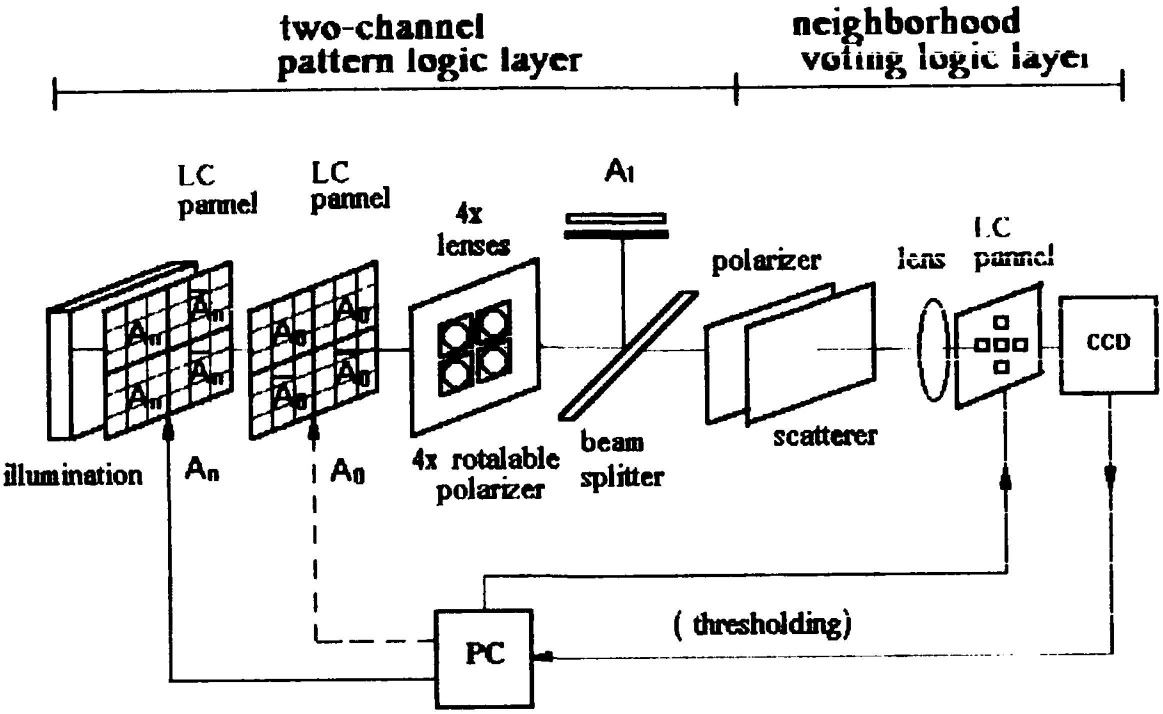
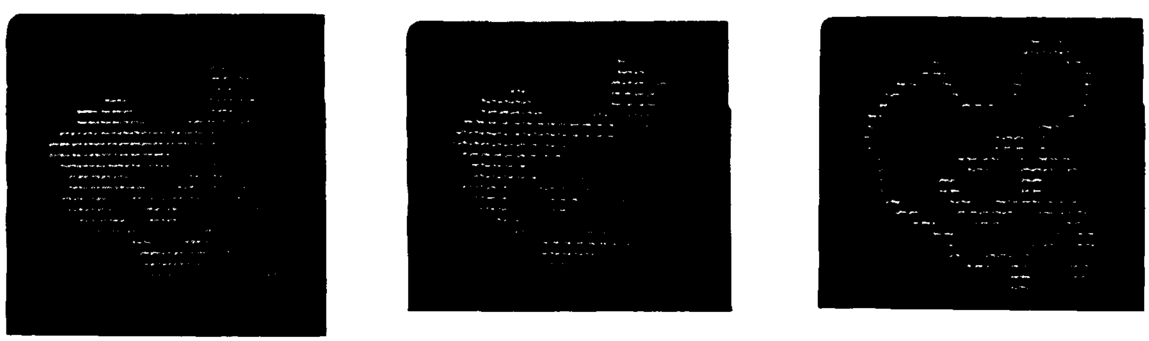


Fig.3. Construction of optoelectronic array



input erosion edge

Fig.4. Processing of a map

Experimental Investigation of the Tolerance of S-SEEDs and L-SEEDs

Frank A. P. Tooley, †Anthony L. Lentine, Sue Wakelin, ‡Alex Wachlowski, Dominic Goodwill, Doug A. Baillie & Kevin Simpson

Physics Dept. Heriot-Watt Univ. Edin. U.K. EH14 4AS
†AT&T Bell Laboratories, Naperville, Illinois
‡Alcatel, Vienna, Austria.
Tel. UK 31 451 3065 FAX 3136

Logic-SEEDs (L-SEEDs)[1] and symmetric-SEEDs (S-SEEDs)[2] have been obtained from AT&T Bell Laboratories. L-SEEDs are a smart pixel technology consisting of electrically connected quantum well p-i-n diodes that can implement complex functions beyond simple NOR gates. Several experiments are currently in progress which will compare the tolerance to spatial power variations of L-SEEDs and S-SEEDs[3]. These include quantifying the effects of the preset[3], differential attenuator[4] and/or offset bias[5] techniques which are used to provide increased tolerance. In addition to presenting experimental results of the operation of these devices as arrays of programmable logic gates, we will present a model with which the optimum operating points for various circuit configurations can be found.

We have recently concluded an experimental study of the characteristics of the S-SEED array. This has involved quantifying the effects of using different voltages, wavelengths and powers. The effect of the saturation of the exciton absorption at high power causes a reduction in contrast from greater than 4:1 at low power to less than 3:1 at 100's microwatts, this effect is illustrated in figure 1. The effect of increasing voltage is an increase in loop width and contrast. A static model has been developed to model these observations.

In addition, we have quantified the effect of small translational misalignments on the effective characteristic. The 3μm diameter beam has been positioned at different locations across the 5μm by 10μm window and the shape of the characteristic noted. This study reveals the precision with which it is necessary to align the system. Alternatively, if the setability of the position of the beams in the windows can be measured, it allows us to quantify the degradation in the characteristic caused by non-ideal alignment.

The required loopwidth and output contrast for correct operation of S-SEED logic gates has been the subject of a study[5] which concludes that loopwidth should be maximized and tolerance is relatively insensitive to output contrast. The predictions of the model were found to agree with experiment. An optical circuit was constructed with which primitive image processing such as noise removal[6] could be performed. The performance of this system when spatial noise was introduced was characterised. An example of the results of this investigation was the measurement of a non-uniformity of 25% and the prediction that with this non-uniformity some of the gates would not operate correctly with the low loopwidth and contrast used. This prediction was confirmed by experiment.

In addition, this optical circuit was characterised for operational errors at different voltages and with defocus introduced. Full details will be presented of this characterisation. It is found that it is impossible to get 100% of the 128 data channels working simultaneously. The belief is that by using L-SEEDs, the tolerance to misalignment and spatial noise will be reduced to an extent that error-free operation will be possible.

The tolerance to spatial noise of S-SEED can be increased using a technique described in ref. [4] and in detail in refs. [3]&[5] which requires differential attenuation of the outputs. Unfortunately, this technique cannot be used if the gate needs to be programmed using the preset[2] to be either a NOR or a NAND gate. The need with the CLIP architecture[7] to have programmable gates has forced us to consider a novel way of increasing the tolerance to spatial noise:- offset bias. This technique consists of simply leaving the preset beam on during the write cycle. One advantage of L-SEEDs may be the lack of the need for this offset bias or of a preset.

The paper will contentrate on recent results which have been obtained with a set-up which uses custom opto-mechanics to ensure that the stability of the system is adequate to obtain reproducible results. A photograph of this set-up is shown as figure 2. It consists of a number of slots in which all the elements are placed. By placing beamsplitters and mirrors at the intersection of slots, a flexible configuration of several independently controllable beams(position and power) can be obtained.

[1] A.L. Lentine, F.A.P. Tooley, et al., "Logic SEEDs: Quantum Well Optoelectronic Multiport Logic Gates, Multiplexers, Demultiplexers and Shift Registers", J. Quant. Electron., 28(6) 1539-1553 (1992); see also A.L. Lentine et al., Appl. Opt., 29, 2153 (1990).

[2] A.L. Lentine, et al., "A 2 Kbit array of S-SEEDs", IEEE Photonic Technol. Lett., 2, 51-53 (1990).

[3] F.B. McCormick, F.A.P. Tooley, et al., "S-SEED-based photonic switching network demonstration", Applied Optics 31(26) 5431-5446 (1992).

[4] N. C. Craft and M. E. Prise, "Optical systems tolerances for S-SEEDs in Optical Computers," Topical Meeting on Optical Computing, Technical Digest Series 1989, volume 9, (Optical Society of America, Washington, DC 1989) pp334-337.

[5] S. Wakelin and F. A. P. Tooley, "The tolerance of differential logic gates to spatial noise", to appear in Optical Computing and Processing, 1992.

[6] F. A. P. Tooley and S. Wakelin, "Design of a S-SEED Cellular Logic Image Processor", to appear in Applied Optics 1992.

[7] R. G. A. Craig, B. S. Wherrett, A. C. Walker, F. A. P. Tooley and S. D. Smith, "Optical Cellular logic image processor- Implementation and programming", Applied Optics 30 2297 (1991).

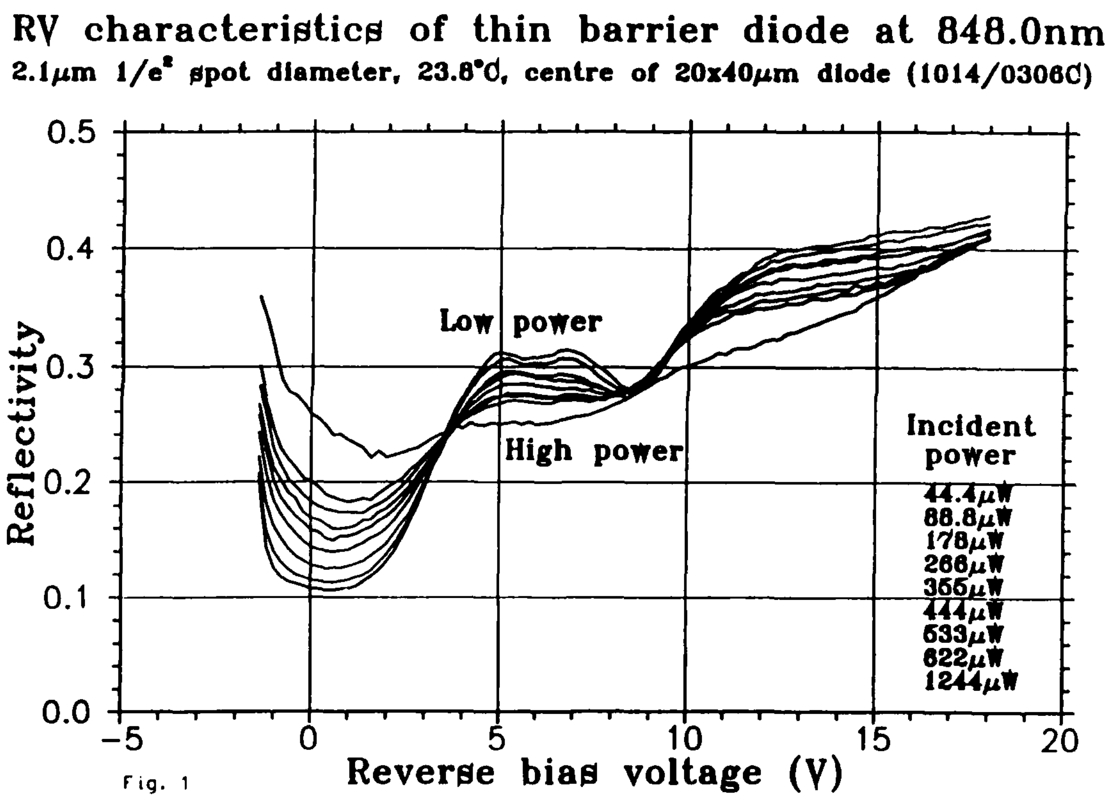
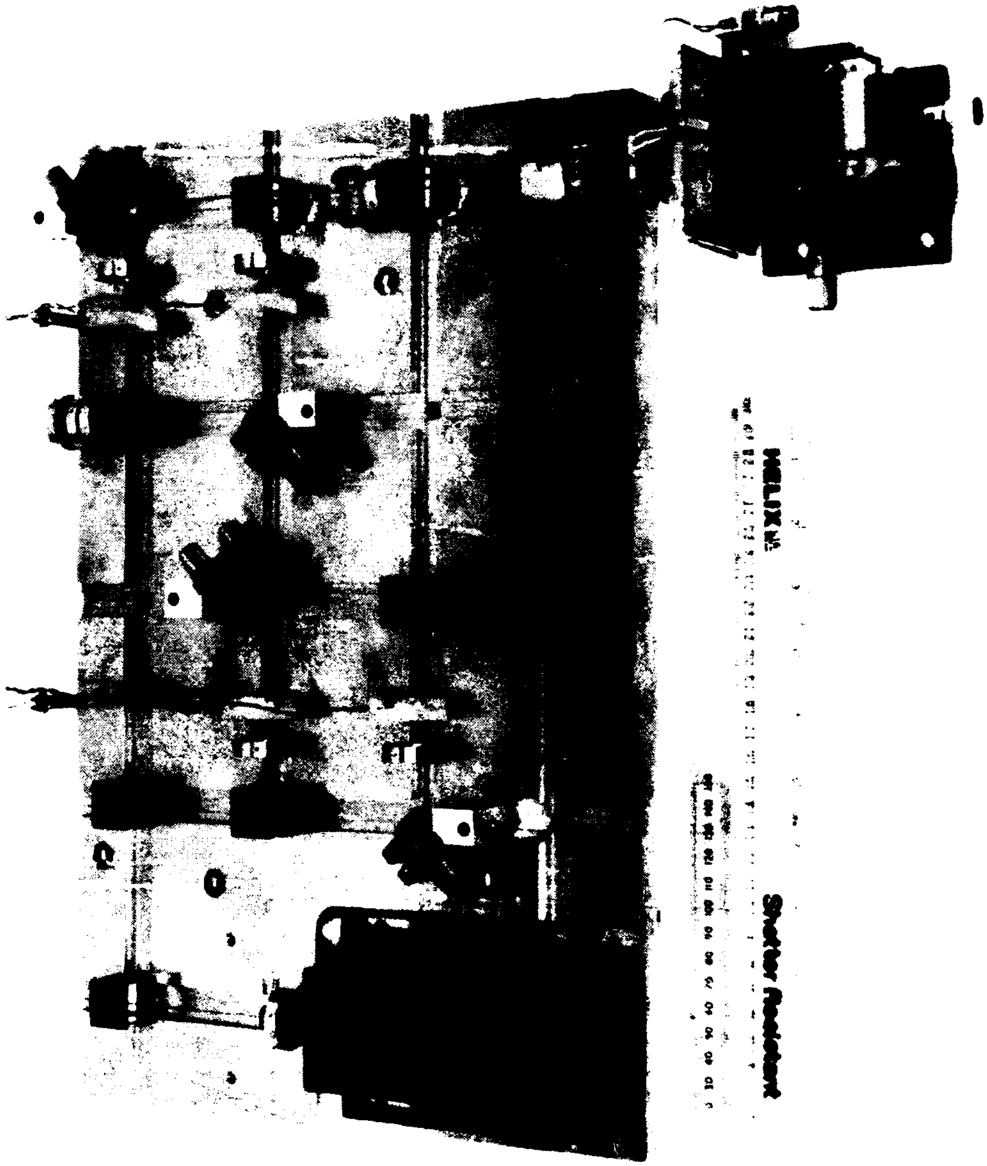


Fig. 1



Paper Withdrawn

A Neural Network Router for Optical Interconnection Networks

C. Lee Giles*and Mark W. Goudreau† NEC Research Institute, Inc.
4 Independence Way, Princeton, N.J. 08540

1 Introduction

Artificial Neural Networks (ANNs) of the type that were described by Hopfield [8] are capable of finding good solutions for certain optimization problems [7]. Furthermore, these ANNs can also solve certain constraint satisfaction problems, which can often be modelled as optimization problems that have numerous correct solutions that are of equal worth. The routing of a set of messages through a multistage interconnection network (MIN) can be modelled as a constraint satisfaction problem. Generating the control bits that define the routes for the MIN can be a time consuming process that can cause a bottleneck in the system.

An ANN has been designed that can potentially solve the problem faster than conventional means. This neural network solution, however, is only applicable to a particular class of interconnection network; one that is constructed out of layers of complete or incomplete crossbar switches. There is no restriction on the connections between successive layers. But there can be no feedback connections and no connections that skip a layer of the interconnection network.

Electronic MINs (EMINs) are usually considered to be 2-dimensional, while optical MINs (OMINs) are generally 3-dimensional. However, the structures of both EMINs and OMINs are such that the routing problem is similar in both cases.

The system being considered will have a set of input ports and a set of output ports that are connected via the MIN. The problem addressed here will be point-to-point communication; no broadcasting will be allowed. Additionally, the circuit switching problem is addressed: for an input port to communicate with an output port, the entire route has to be established and maintained for a certain period of time. At the beginning of each message cycle, an input port can decide that it wants to communicate with an output port. So for each message cycle there is a set of desired messages. The problem is to generate control bits for the MIN given the message set.

Figure 1 contains a diagram of a communication system with a neural network router. The neural network router is discussed in [3, 4]. In these references the neural network router was applied to an EMIN routing problem. The very same methodology, however, can be applied to a 3-dimensional OMIN problem by collapsing the 3-dimensional problem down to a 2-dimensional problem. Since Hopfield model neural networks can be constructed efficiently using an optical system (for example, see [1]), the neural network routing methodology is particularly appealing for an optical computing environment.

The neural network router in Figure 1 has two logic blocks to interface the neural network with the rest of the communication system. The logic block *Logic1* converts the desired message set into a language the neural network can understand, namely bias currents. The logic block *Logic2* is required to convert the routing array solution of the neural network into a crosspoint form that the interconnection network can understand.

Previous work on utilizing Hopfield model neural networks to facilitate communication through interconnection networks can be found in the collection of references given in [4]. In this paper, we propose a neural network router which may be optically implemented to control an optical interconnection network.

2 Optical Networks

There has been much interest in using optical technology for implementing interconnection networks and switches; for a collection of such work see [2]. Our concern will be on optical multistage interconnection networks [6, 9, 10, 11, 12, 14]

Consider the interconnection network in Figure 3. It is an example of a 16 × 16, 3 stage OMIN. The basic building block is a 4 × 4 optical crossbar switch [5, 13]. The OMIN has 12 of these switches. Each crossbar switch in the OMIN is connected to each crossbar switch in the neighboring stages. This MIN is capable of routing any input/output permutation.

It is straightforward to map the 3-dimensional routing problem to 2-dimensions. Figure 4 is a 2-dimensional representation of the 3-dimensional MIN from Figure 3.

3 The Routing Representation

A routing representation for routes in a MIN will now be constructed. This representation will be called the

*also UMIACS, University of Maryland
†also Electrical Engineering, Princeton University

routing array, and it is the same representation that is used by the ANN. In Section 4, the appropriate ANN structure will be described. The neural network will be in a state of minimal energy when the neuron outputs directly represent a legal routing array.

Every message route that can be established through a MIN has a corresponding routing matrix. The columns of a routing matrix represent the stages of the interconnection network, while the rows represent the output ports for each stage of the interconnection network. If $a_{i,j} = 1$, the message is routed through output port i of stage j. Having $a_{i,j} = 0$ implies that the message is *not* routed through output port i of stage j. As an example, any routing matrix for the MIN shown in Figure 4 will be a 16×3 matrix. Every element in the routing matrix will either be a "0" or a "1".

The routing array for the set of messages is simply constructed by treating each routing matrix as a "slice" and constructing a "loaf". The routing array is a 3-dimensional representation of a set of routes, and each slice of the array represents a single route. For our example, there are two messages to be routed so the routing array will have two slices. In general, if a system has m input ports, as in Figure 1, there can be m slices in the routing array.

Each element of the routing array now has three indices. If element $a_{i,j,k}$ is equal to 1 then message i is routed through output port k of stage j. We say $a_{i,j,k}$ and $a_{l,m,n}$ are in the same *row* if $i = l$ and $k = n$. They are in the same *column* if $i = l$ and $j = m$. Finally, they are in the same *rod* if $j = m$ and $k = n$.

A legal routing array will satisfy the following three constraints: (1) one and only one element in each column equal to 1; (2) the elements in successive columns equal to 1 represent output ports that can be connected in the interconnection network; and (3) no more than one element in each rod equal to 1.

The first restriction ensures that each message will be routed through one and only one output port at each stage of the interconnection network. The second restriction guarantees that each message will be routed through a legal path in the interconnection network. The third restriction resolves any resource contention in the interconnection network. In other words, only one message can use a certain output port at a certain stage in the interconnection network. When all three of these constraints are met, the routing array will provide a legal route for each message in the message set.

4 The Neural Network Router

The construction of an ANN in which each neuron directly represents an element in the routing array for an interconnection network and message set is a three-dimensional structure just like the routing array. Each $a_{i,j,k}$ of a routing array is represented by the output voltage of a neuron, $V_{i,j,k}$. At the beginning of a message cycle, the neurons have a random output voltage. If the ANN settles in one of the global minima, the problem will have been solved.

A synchronous Hopfield model neural network is used [7, 8]. The value of τ, from [7], was set to 1.

The ANN is forced into stable states that are the local minima of the energy equation:

$$E = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} T_{ij} V_i V_j - \sum_{i=1}^{N} V_i I_i \qquad (1)$$

Now an energy function is constructed such that the neural network will be in a global minima when it directly represents a legal routing array. The derivation of the correct energy function is too lengthy to be shown here; a detailed derivation is given in [3, 4].

The end result is that the connection and bias current values from Equation 1 are shown here:

$$\begin{aligned} T_{(m1,s1,p1),(m2,s2,p2)} = & \qquad (2)\\ & -(A+C)\delta_{m1,m2}\delta_{s1,s2}(1-\delta_{p1,p2}) \\ & -B\delta_{s1,s2}\delta_{p1,p2}(1-\delta_{m1,m2}) \\ & -D\delta_{m1,m2}[\delta_{s1+1,s2}d(s2,p1,p2) \\ & +\delta_{s1,s2+1}d(s1,p2,p1)] \end{aligned}$$

$$\begin{aligned} I_{m,s,p} = & \qquad (3)\\ & C - D[\delta_{s,1}d(1,\alpha_m,p) + \delta_{s,S-1}d(S,p,\beta_m)] \end{aligned}$$

The function $\delta_{i,j}$ is a Kronecker delta ($\delta_{i,j} = 1$ when $i = j$, and 0 otherwise). A, B, C, and D are arbitrary positive constants. The function $d(s1, p1, p2)$ represents the "distance" between output port $p1$ from stage $s1 - 1$ and output port $p2$ from stage $s1$. If $p1$ can connect to $p2$ through stage $s1$, then this distance can be set to zero. If $p1$ and $p2$ are not connected through stage $s1$, then the distance can be set to one. The function α_m is the source address of message m, while β_m is the destination address of message m.

The connection values from Equation 2 are well defined for a given MIN. That is, once the MIN is designed, the neural network and all of its inter-neuron connections can be calculated. When different groups of input-output ports need to be connected, it is only the input bias currents of boundary neurons that are affected. Equation 3 quantifies that change.

If the user has the ability to make the output of a rod of neurons equal to zero or give a rod of neurons a large negative input bias current, then the neural network can provide a fault-tolerant routing scheme. For example, if an output port in some stage of the interconnection network is faulty, the user could set the rod of neurons that represents that output port to zero. The rest of the neural network can operate exactly as it did before. Neither the structure nor the weights need to be changed. Similarly, this routing methodology could tolerate faulty input ports and broken buses. However, the user must know if a fault exists and where it exists in the MIN.

One major drawback of the neural network routing methodology is that it requires a large number of processors (neurons) to generate a solution. For example, $O(M^2S)$ processors are required if there are S stages in the interconnection network, M inputs, M outputs, and M connections between two consecutive stages.

5 Simulation Results

A synchronous ANN was simulated and the results are shown in Figure 5 for routing in the OMIN shown in Figure 3. For the OMIN shown in Figure 3, the neural network router contained $16 \times 16 \times 2 = 512$ neurons. Each data point in the graph gives the results for a set of 1000 message cycles. The x-axis represents the number of messages that are to be routed in each message cycle. The message pairs are generated randomly with no input port or output port conflicts allowed. The y-axis represents the expected number of messages routed in each message cycle. A message is considered to have been routed successfully when the neural network provides it with a legal routing and when the message has no contention with any other message for an output port on in any stage of the MIN.

For the OMIN shown in Figure 3, any I/O permutation is possible. Thus, any message that was not routed correctly implied a failure in the neural network solution, not a limitation in the MIN.

For the purposes of the simulations, A, C, and D from Equations 2 and 3 were all set to 3.0. B was set to 6.0. These parameters were chosen experimentally, and the performance of the neural network router is highly dependent upon them. The analog neural network was simulated with a digital computer.

The simulations show that increasing the number of messages degrades the performance of the system. Other simulations show that the performance also degenerates when the MIN becomes more complicated.

6 Conclusions

A neural network routing methodology was presented that is capable of providing control bits to optical multistage interconnection network (OMINs). It was shown how the 3-dimensional OMIN can be reduced to a 2-dimensional MIN, making the neural network routing solution described in [3, 4] possible. This routing method is valid for a wide range of OMINs that have a multistage structure. It was shown that the routing method is fault-tolerant. Once the OMIN is chosen, the routing neural network can be constructed and the weights never have to change. For a new message set, only certain boundary bias currents are changed.

The usefulness of this routing method depends on the speed of the implementation of Hopfield neural network as well as other requirements of the system. The fact that a Hopfield neural network can be readily constructed in an optical computing environment makes the neural network routing approach quite attractive for OMIN routing problems.

For our simulations, the routing performance degraded as the MIN size increases and as the number of messages in a message cycle increases. Preliminary results show that the performance of the neural network router will fall off considerably as the number of stages in an MIN increase. However, such results are dependent on neural network parameters. Depending on the implementation of the neural network router, the routing method described in this paper might have advantages over many other routing methods in terms of speed. Furthermore, the neural network routing methodology can be applied to many irregular OMINs that have no deterministic routing scheme that is any better than an exhaustive search. Thus, it is felt that the neural network routing methodology is most suitable for establishing routes for irregular OMINs and OMINs that do not have self-routing capabilities. While no optical implementation of the neural network router is explicitly proposed, such a neural network could be implemented optically.

References

[1] N. Farhat, D. Psaltis, A. Prata, and E. Paek, *Applied Optics*, vol. 24, pp. 1469–1475, 1985.

[2] J. Goodman and L. Hutcheson, eds., *Applied Optics, Information Processing, Special Issue on Optical Interconnects*, vol. 29. Optical Society of America, March 1990.

[3] M. Goudreau and C. Giles, in *Advances in Neural Information Processing Systems 4* (J. Moody, S. Hanson, and R. Lippmann, eds.), (San Mateo, CA), pp. 722–729, Morgan Kaufmann Publishers, 1992.

[4] M. Goudreau and C. Giles, *International Journal of Neural Systems*, vol. 3, no. 2, 1992.

[5] A. Himeno and M. Kobayashi, *Journal of Lightwave Technology*, vol. LT-3, pp. 230–235, April 1985.

[6] H. Hinton, *IEEE Communications Magazine*, vol. 28, no. 4, pp. 71–89, 1989.

[7] J. Hopfield, *Biological Cybernetics*, vol. 52, pp. 141–152, 1985.

[8] J. Hopfield, *Proceedings of the National Academy of Science USA*, vol. 81, pp. 3088–3092, May 1984.

[9] S. Kawai, in *Photonic Switching II* (K. Tada and H. Hinton, eds.), pp. 216–219, Berlin: Springer-Verlag, 1990.

[10] F. Kiamilev, P. Marchand, A. Krishnamoothy, S. Esener, and S. H. Lee, *IEEE Journal of Lightwave Technology*, to be published.

[11] M. J. Murdocca and T. J. Cloonan, *Applied Optics*, vol. 28, no. 13, pp. 2505–2517, July 1989.

[12] L. Rudolph, D. Feitelson, and E. Schenfeld, in *Optical Interconnections and Networks*, vol. 1281, SPIE, March 1990.

[13] A. Sawchuck, B. Jenkins, C. Raghavendra, and A. Varma, *Computer*, vol. 20, no. 6, pp. 50–60, June 1987.

[14] C. Stirk and R. Athale, *Applied Optics*, vol. 27, p. 1721, May 1988.

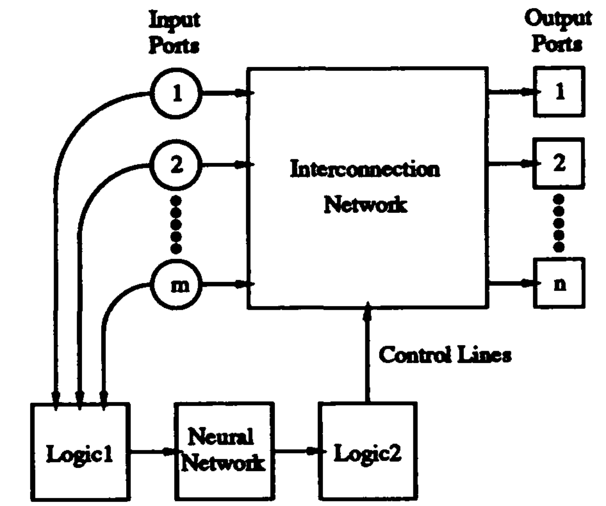


Figure 1: The communication system with a neural network router.

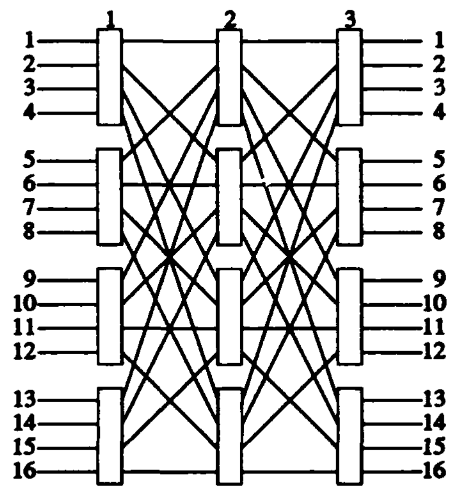


Figure 4: A 2-dimensional representation of the interconnection network shown in Figure 3.

Interconnection Network Implementation	*Neural Network Router Implementation*	
	Electrical	*Optical*
Electrical	*Implemented*	—
Optical	*Implied*	*Proposed*

Figure 2: Neural network router implementations and interconnection network implementations.

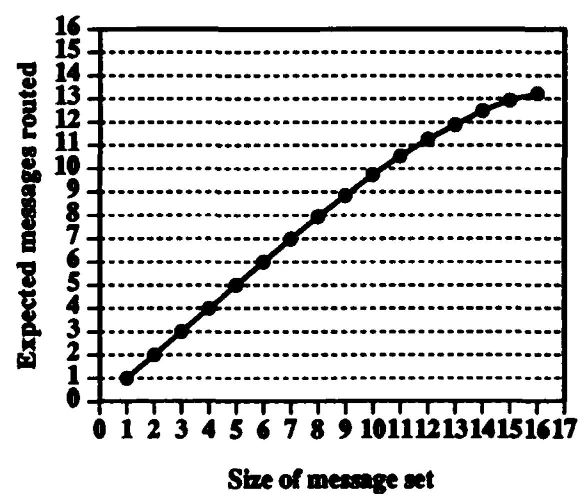


Figure 5: Routing results for the interconnection network shown in Figure 3.

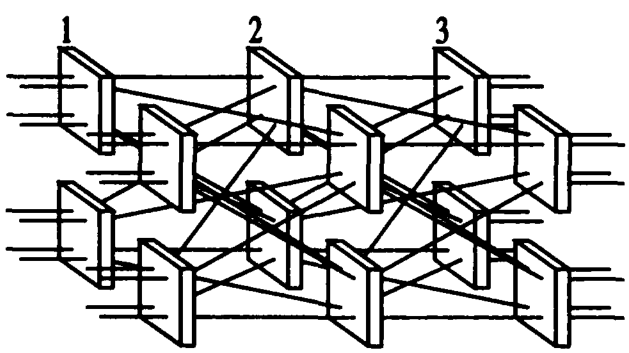


Figure 3: A 3-dimensional optical interconnection network using 12 4x4 crossbar switches.

Design of a microlens based total interconnection for optical neural networks

P.C.H. Poon, D.R. Selviah and J.E. Midwinter
Department of Electronic and Electrical Engineering, University College London, Torrington Place, London WC1E 7JE, UK (Tel) +44 71 387 7050 (Fax) +44 71 387 4350 (e-mail) d.selviah@ uk.ac.ucl.eleceng

D. Daly
National Physical Laboratory, Teddington, Middlesex, UK TW11 0LW (Tel) +44 819773222 (Fax) +44 819432155

M.G. Robinson
Sharp Laboratories of Europe Ltd., Edmund Halley Road, Oxford Science Park, Oxford, UK OX4 4GA (Tel) +44 865747711 (Fax) +44 865747719

1. INTRODUCTION

The ability of neural networks to quickly adapt to changing circumstances is important for applications ranging from fast recognition of images to adaptive line equalisation for telecommunication systems. Optics offers the ability to realise numerous parallel, wide bandwidth, interconnections with low crosstalk. In this paper we investigate the practical design of an optical total interconnection between two planar arrays of optical elements for use in an optical neural network. Each pixel in the first array of modulators or surface emitting LEDs or lasers is to be connected to every pixel in the second array of detectors and non-linear thresholding elements. The rapidly programmable, analogue weights required can be imposed on the optical connections by an electrically addressed, liquid crystal, nematic Spatial Light Modulator (SLM) placed between the arrays. If the pixels are small enough the light from each pixel in the first plane will diffract onto every pixel in the second plane performing the required interconnection. This simplistic approach suffers from two drawbacks, the scattering distribution is not isotropic from each pixel and much of the light is diffracted away from the second array and so results in high loss. A variant uses a Fourier Transform lens to reduce the length of the system but suffers from the same problems. A second approach is to use an array of Fresnel near-field holograms between the two planes each redirecting the light from one pixel in the first plane towards one in the second plane and also focussing the light onto the detector pixel reducing the loss. However, redirection of light through large angles requires fine resolution features in the hologram and also is rather lossy when thin holograms are used. A third approach is to use a Dammann far field hologram, but this requires additional elements, the Fourier transform lenses, which must also be aligned and for which space must be found. A fourth approach is to place a two dimensional array of refractive microlenses between the two planes. This approach is used in this paper and has the advantage that such arrays are commercially available [1].

2. MICROLENS BASED TOTAL INTERCONNECTION SYSTEM DESIGNS

In the "image multiplexer" total interconnection design (Fig. 1) [2] each microlens in an N^2 element array is used to image the whole of the NxN input plane to the output which, ideally, has N^4 elements. The SLM is placed next to the output plane to control each of the N^4 connection weights independently. However, it is difficult to realise such a high density of elements on the output plane as it will have similar dimensions to the input array and microlens array due to the short microlens focal lengths. In the "shared microlens" total interconnection system (Fig. 2) [3] an array of $(2N-1)^2$ microlenses images each of the N^2 input pixels onto each output pixel. An individual microlens connects several inputs and outputs. A similar interconnection system [4] was demonstrated using pinholes rather than microlenses. The connection weights would be imposed in by placing an SLM at the microlens array and only independently weighting groups of interconnections passing through each microlens. Clearly, this would reduce the number of degrees of freedom. We have used the theoretical analysis of reference [3] as a basis for practical system design. An important practical parameter is the system length, L,which has upper and lower limits caused by diffraction and off-axis aberrations respectively. If the input and output planes are close to the microlens array, then the off-axis angles for the furthest elements at the corners of the planes are large. One of the most serious aberrations in this case is astigmatism which increases quadratically with distance from the microlens axis and which cannot easily be compensated. The upper limit to the system length is determined by the size of the diffraction limited spot on the output element which increases with system length and depends on the focal length, f of the microlens. The limit is assumed to be when the Airy disc is the same size as the pixel on the output plane and also assumes that crosstalk due to the 16% of the focal energy in the sidelobes which fall on adjacent pixels is acceptable. The upper and lower limits for system length, L, are given below for the image multiplexer and shared microlens system respectively

$$\frac{d^2}{2.44\lambda} \geq L \geq 2N^{\frac{3}{2}} d I^{\frac{1}{2}}$$

$$\frac{d^2}{2.44\lambda} \geq L \geq 2N d I^{\frac{1}{2}}$$

where d is the microlens pitch in the array and I is the dimensionless strength of the astigmatic aberration. If the microlenses are thin and unmounted on a substrate then I is given by $(3n+1)/4n$ where n is the refractive index of the photoresist [3]. The diffraction limit is the same for both systems and is plotted as a parabolic function of d in figures 3 and 4. The astigmatic limit appears as straight lines on these plots of system length against microlens pitch; the slope of the lines increases as the number of elements in the array increases and as the astigmatic aberrations in the microlens worsen. For example, if $N = 32$, then a practical system design would be possible in the shaded regions in the figures to satisfy the inequalities. The figures show that more elements can be connected in the shared thin microlens system using a smaller microlens pitch and in an overall more compact system.

3. EXPERIMENTAL MICROLENS EVALUATION

The two most promising types of microlens array for use in the total interconnection system are those made by ion-exchange [1] and those made by melting photoresist [5-8]. Focal lengths of about 70-1300µm with numerical apertures of about 0.2-0.5 and lens diameters of about 5-1000µm are possible. However, more detailed characteristics are required in order to correctly design the system. We characterised 4 commercially available samples of 32x32 ion-exchanged microlens arrays of pitch 250µm and focal length 560µm. We also fabricated our own two dimensional arrays of photoresist microlenses using the following technique. We spun Hoechst AZ 4620-A photoresist onto a glass microscope slide at a speed of 4500rpm giving a resist thickness of 6µm. An electon beam fabricated mask of an array of 64x64 circles of 120µm diameter and pitch 125µm was used lithographically to form similarly shaped circular islands of photoresist on the glass substrate. The photoresist was melted at a temperature of 210°C on a hot plate for 3minutes and left to cool. During this time the surface tension pulled the surface into the required convex shape for operation as a microlens. The focal length was found using a microscope to be 300μm giving a numerical aperture of 0.20 and an f-number of 2.4 in a range also studied in references [6] and [7]. The complete characterisation of the wavefront aberrations was performed by an interferometric technique [8] for both of the microlens arrays. Only the results for the photoresist microlenses are shown in Figure 5. A Zernike polynomial fit to the wavefront (Fig 5(a)) was made and was used to calculate the primary Seidel aberration coefficients. The ideal diffraction limited focal spot width for this lens size and focal length using 632.8nm wavelength would be 1.9μm between the first nulls and that found by calculation from the wavefront aberration (Fig 5(b)) is 3μm. The 3dB width is found to be 1.6μm. The Strehl ratio was 0.75 for the photoresist microlenses (i.e. 0.67 wavelength aberration variation in the wavefront, mainly spherical) whereas that for the ion-exchanged lenses was 0.05 (i.e. a variation of 2.22 wavelengths in the wavefront, mainly spherical). This difference was also clearly apparent in the point spread functions and in the modulation transfer functions. The primary Seidel astigmatic aberration coefficient was 25% larger, the comatic aberration coefficient was 82% larger and the spherical aberration coefficient was 135% larger for the ion-exchanged microlenses than for the photoresist lenses. The ion-exchanged lens numerical aperture, 0.22 was comparable with that of the photoresist lenses, 0.20 but the aberration differences observed may not persist as the focal lengths and diameters of each lens are scaled up and down.

4 CONCLUSIONS

Two microlens total interconnection systems and two types of microlens array were evaluated for use in an optical neural network and results suggest that the optimum microlens total interconnection system would be one employing photoresist microlens arrays in a shared microlens configuration. Such a system has yet to be demonstrated. The authors thank Mike Hutley of the National Physical Laboratory, UK for use of the interferometric measurement system and Sharp Laboratories of Europe Ltd., UK for financial support of one of the authors, P. C. H. Poon.

REFERENCES

1. K. Nishizawa and M. Oikawa, "Planar microlens by ion-exchange", Inst. of Phys. Meeting on Microlens Arrays Proc., ed. M. C. Hutley, No. 30, 17-21 (1991)
2. A. Akiba and K. Iga, "Image multiplexer using a planar microlens array", *Appl Opt*, **29**(8),1094-1100 (1990)
3. T. Sakano, K. Noguchi and T. Matsumoto, "Optical limits for spatial interconnection networks using 2-D optical array devices", *Appl. Opt.*, **29**(8), 1094-1100 (1990)
4. K. Noguchi and T. Sakano, "Optically implemented Hopfield associative memory using two-dimensional incoherent optical array devices", IJCNN-90, Washington D.C. II-68 -II-71 (1990)
5. Z.D. Popovic, R.A. Sprague and G.A.N. Connell, "Technique for monolithic fabrication of microlens arrays", *Appl. Opt.*, **27**(7), 1281-1284 (1988)
6. A.Y. Feldblum, C.R. Nijander, W.P. Townsend and C.M. Mayer-Costa, "Performance and measurements of refractive microlens arrays", *SPIE Proc.* **1544**, 200-208 (1991)
7. D. Daly, R.F. Stevens and M.C. Hutley, "The manufacture of microlenses by melting photoresist", Inst. of Phys. Meeting on Microlens Arrays Proc., ed. M. C. Hutley, No. 30, 23-34 (1991)
8. M.C. Hutley, D. Daly and R.F. Stevens, "The testing of microlens arrays", Inst. of Phys. Meeting on Microlens Arrays Proc., ed. M. C. Hutley, No. 30, 67-81 (1991)

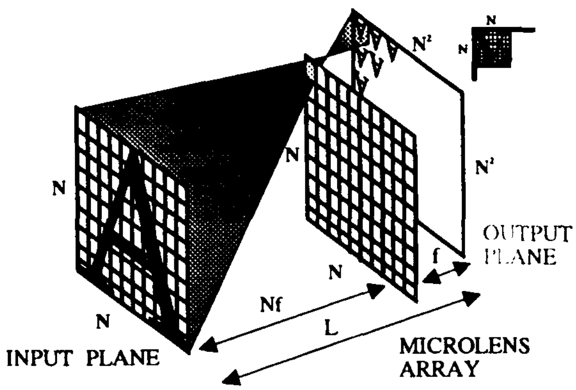


Fig.1 Image multiplexer total interconnect

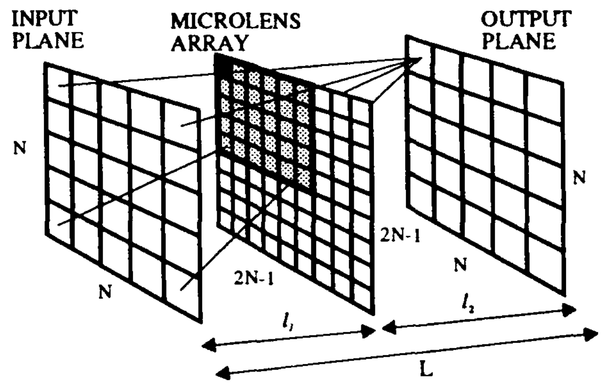


Fig. 2 Shared microlens total interconnect

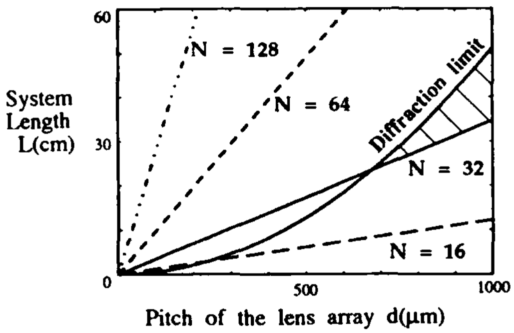


Fig. 3 Dependence of system length on microlens pitch in image multiplexer

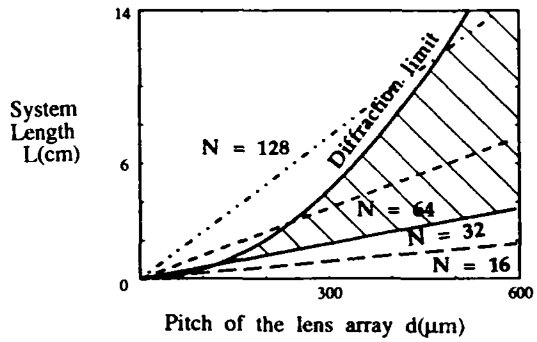
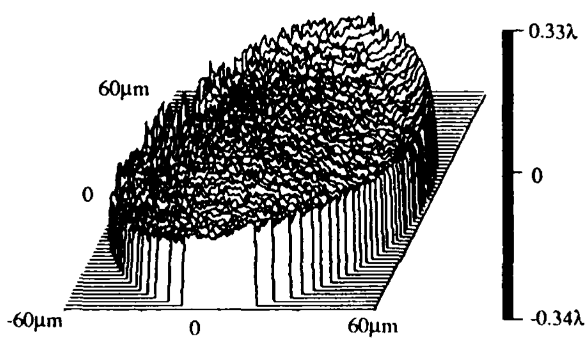
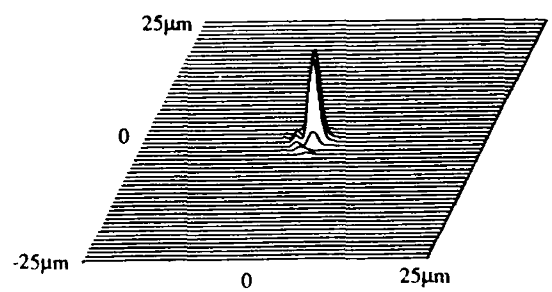


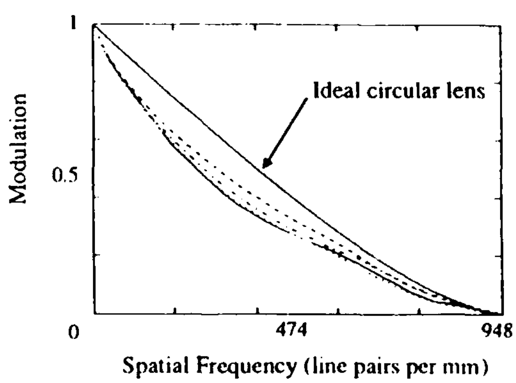
Fig. 4 Dependence of system length on microlens pitch in shared microlens system



(a)



(b)



(c)

Fig. 5 Photoresist microlens experimental characterisation; ϕ 120 μm, f = 300 μm

(a) Wavefront aberrations

(b) Point spread function

(c) Modulation transfer function: Curves at four 45° angled cross sections

High-Speed Database Processing on an Optical Content-Addressable Parallel Processor(OCAPP)

Ahmed Louri and James A Hatch Jr.

Department of Electrical and Computer Engineering
The University of Arizona
Tucson, AZ 85721
Phone : (602) 621-2318
E-mail : louri@paris.ece.arizona.edu

Introduction

Recently, an efficient optical architecture, known as Optical Content-Addressable Parallel Processor(OCAPP), has been introduced for the efficient implementation of symbolic computing tasks [1]. OCAPP simultaneously performs multiple equivalence and threshold searches {greater than(G), less than(L), greater than or equal(GE), less than or equal(LE)} in addition to compound operations such as sorting. Architecturally, OCAPP consists of a *selection unit*, a *match/compare unit*(MCU), a *response unit*, a *control unit*, and an *output unit*, see Figure 1 and ref [1] for more details. The selection unit holds n words, where each word is m-bits long, and contains word and bit-slice selection logic to control which bits participate in comparisons. The MCU is the central element of OCAPP since it performs parallel comparisons between the words of the $n \times m$ storage array and the $1 \times m$ interrogation register. The results are reported in three registers, R, G, and L, corresponding to the equals, greater than, and less than operators, respectively. The output unit transfers words to the output plane which becomes an input of the host system requesting the operation.

Database processing is based upon the manipulation of tables, known as relations. A relation is a file similar to ones that contain information about employees in an employee database. The rows of a relation, known as tuples, are data records about the subject of the relation, such as a single employee. Referred to as attributes, the data fields of a tuple are the columns of a relation. Examples of attributes in the employee database are employee name and ID number. A set of operators known as the *relational algebra* is responsible for the manipulation of relations and is composed of *union, intersection, difference, selection, projection,* and *join.* The relational operations operate on one or two relations to produce an output relation. Union, intersection, and difference require equivalence searching of entire tuples while selection, projection, and join require the equivalence and/or threshold searching of specific attributes. Since relational operations are applied to the entire table, e.g. several tuples simultaneously, they are well suited for parallel execution on OCAPP. Due to their inherent parallel nature, relational operations will realize an increase in throughput when implemented on OCAPP. Improvements in database algorithms are realized through implementations that further exploit this inherent parallelism.

Extension of OCAPP for High-Speed Database Processing

The OCAPP Database Machine contains a *selection unit*, a *2-D match/compare unit*(2-D MCU), an *equality unit*, a *threshold unit*, a *relational algebra unit* and a *deflection/output unit.* See Figure 3 for details. The equality unit processes the R register used extensively in union, intersection, and difference. The threshold unit is responsible for constructing the G, L, GE, and LE registers. It is a rather complex unit due to the use of an improved thresholding algorithm that requires only a single iteration to perform t thresholds, whereas current algorithms require s iterations, t being the number of tuples in a relation and s being the number of columns in an attribute. Hence, the algorithm will realize a speedup factor of s. The relational algebra unit processes the results in the R, G, L, GE, LE registers in accordance with the rules of the current relational operation to supply the deflection/output unit with the proper electrical control signals. Deflection in an output unit performs the parallel transfer of words from Relations A and B to their respective locations in the output, Relation C. This is important because certain algorithms transfer only a subset of Relations A or B to Relation C. Since the tuples in the output must appear in consecutive rows, a method of mapping say the third tuple of Relation A to the second row of Relation C is necessary. Traditionally, this mapping is accomplished through the iterative transfer of tuples of Relations A and B to Relation C after a comparison is made and the registers are written. However, this reduces the parallelism of the machine. By transferring multiple rows to the output plane simultaneously, this is no longer a limitation.

The selection unit contains an $n \times m$ storage array and a dynamic transmissive mask that disables bits by blocking the light that would propagate to the 2-D MCU. Composed of a 2-D Optical Content-Addressable Memory(OCAM), the 2-D MCU is implemented with the matrix-matrix multiplier presented by Gheen [2] and shown in Figure 2. The advantages of a 2-D OCAM are obvious. By performing matrix-matrix comparisons, the throughput is effectively increased by a factor of n. The matrix-matrix multiplier is a Stanford vector-matrix multiplier with a linear phase mask that separates rows of the input matrix (spatial light modulator 1, SLM1) into side-by-side columns in the output plane. The 2-D OCAM reduces to a 1-D OCAM when the off-axis rows of the input matrix are disabled. In order to access the 1-D data, a beamsplitter is inserted between SLM2 and the spherical-cylindrical lens combination (S/CL1). As seen in Figure 2, the 2-D MCU produces the R registers and the unprocessed result of a 1-D vector-matrix comparison. The equality unit contains a matrix CCD that detects the R registers from the 2-D MCU and supplies them in electrical form to the relational algebra unit.

From Figure 4, the on-axis tuple of Relation A, which will be referred to as Tuple A, is copied into each row of SLM3 in the threshold unit. This allows the threshold unit to compare the unprocessed vector-matrix data from the 2-D MCU to the comparand, Tuple A. SLM3 is followed by a holographic interconnect composed of a matrix of sub-holograms, one per bit, that projects in the image plane, each bit onto those to the right of it. Following the hologram is an optical system that provides spatial filtering for the removal of the zero-order and conjugate beams in addition to collimating the overlapping diffracted beams. In Figure 4, the black extension of the optical system is a stop to eliminate the portion of beams that expand past the extent of the bit locations. SLM4 compares the hologram output with a copy of SLM3 output which is then focused to a line with a cylindrical lens and detected with a CCD to form the G register. Throughout the threshold unit, there are four polarizers that eliminate the bits that evaluate to greater than or less than. The deflection/output unit uses a dynamic transmissive mask and electrically controlled deflecting optics to individually manipulate rows onto the output relation.

Implementation of Relational Operations on OCAPP

Due to the short length of this summary, we only discuss a single algorithm to illustrate the use of OCAPP for database processing. We chose the Selection algorithm because it utilizes the major components of the system. The algorithm selects all of the tuples in a relation for which a criteria is met in a particular attribute.

Theta represents the criteria and could be any operator of the set: (greater than, less than, greater than or equal, less than or equal). As an example, we use the selection of employee age in our database where we are searching for all employees older than $6(0110_2)$. See Figures 3 and 4 for more details of the following discussion. Selection is performed by loading 0110_2 into Tuple A which corresponds to the on-axis row of SLM1 while instructing the selection unit to disable the other input tuples. Tuple A is also loaded into each row of SLM3 in the threshold unit. The attribute field upon which selection is performed, the collection of employee ages in this case, is loaded into SLM2 of the 2-D MCU. The 2-D MCU compares Tuple A and Relation B and writes the R register. It is important to note that the data being output from the 2-D MCU is vertically inverted(upside down). In Figures 3 and 4, the data is kept this way in order to be physically correct. However, the equality and thresholds units are responsible for re-inverting the data so that the system input relations and output relation may be compared with ease.

The single-iteration threshold unit operates on the principle that the output of a vector-matrix comparison before passing through the final cylindrical lens, S/CL1, is a matrix of information such that each bit represents equality (denoted by a logical 0) or inequality (denoted by a logical 1). In our polarization-encoding scheme, a '0' is represented by vertically polarized light while a '1' is represented by horizontally polarized light. If a bit is dark, then it behaves as a don't care. Elimination of undesired bits is accomplished through the use of a polarizer. If the bit-wise comparison evaluates to not equal, it is known that a mismatch occurred, but it is not readily known whether the bit from Tuple A was greater than the corresponding bit from Relation B or vice versa. However, this is resolved through an additional comparison of the bits that evaluated to not equal with the original Tuple A. This comparison is performed in SLM3 of the threshold unit and may be interpreted as follows. First, the bits that evaluated to equality are eliminated, accomplished through the use of horizontal polarizer, P1. If two bits, A_i and B_i, are not equal and the corresponding bit in Tuple A is '0', the output of the comparison is '1'. This reveals on a bit-by-bit basis whether $B_i > A_i$ or $A_i > B_i$ by the presence of a '1' or '0' respectively.

Normally, one would iterate through the bits beginning with the most significant bit (MSB), in order to find the first bit position in which a mismatch occurred. The bit positions below this have no effect and are ignored. Hence, only one bit position is responsible for determining if a word is greater than or less than another, the goal being to determine this. To perform this search in parallel , we place a vertical polarizer, P2, in the path to eliminate the greater than bits, i.e. make them dark. Each bit is projected, through the use of a holographic interconnect, across each of the bits below it in order to disable them if a mismatch occurs. The disabling, which will occur in SLM4, occurs because each of the bits that evaluated to less than will cause a polarization rotation. In addition to eliminating the greater than bits, the less than bits are eliminated in a copy of SLM3 output simultaneously through the use of polarizer, P3. Thus, the input to SLM4 is two data planes, one such that the greater than bits are eliminated, the other such that the less than bits are eliminated. The less than bits cause a ninety degree rotation of a greater than bit and turn it into a less than bit. Less than bits are then eliminated with polarizer P4. If any bits in a row are still bright, then that attribute field of Relation B is greater than Tuple A. The L register may be determined in a similar manner or may be computed from the G and R registers.

The results of the R, G, L, GE, LE registers are then sent to the relational algebra unit which then supplies the necessary controls to the deflection/output unit so that it may map tuples onto Relation C in parallel. More details of the architecture will given at the meeting.

Unfortunately, a single NOR array does not support a 2-D single-iteration threshold operation because it cannot perform multiple NOR operations in a single X-Y location. Since the multiple beams in a 2-D system propagate through SLM3 at various angles, multiple single-interation thresholds can be performed in parallel if a NOR matrix and a holographic interconnect are included for each desired row of the input. In the Selection operation, this is used to perform multiple-argument selections in one iteration because multiple *theta* operations can be evaluated simultaneously. This is also very useful for performing extremum searches (max, min) in a single iteration since the maximum-valued word corresponds to the one with a GE register loaded with ones. This would ultimately be used to create sorting algorithms that only require n iterations since a new maximum-valued word is found in each iteration, as opposed to conventional optical algorithms which require $n \times m$ iterations.

Because the execution time of the Selection algorithm is independent of, l, the number of bits in an attribute, a speedup factor of l can be realized over current optical and electrical Selection algorithms when *theta* is an operation involving $(<, >, \geq, \leq)$. Since Selection based on thresholding is a highly utilized function in database processing, this translates into a substantial increase in overall system performance.

References

[1] Ahmed Louri , "An Optical Content-Addressable Parallel Processor: Architecture, Algorithms, and Design Concepts," *Applied Optics*, vol. 31, no. 17, pp. 3241 - 3258, June 1992

[2] Gregory Gheen , "Optical matrix-matrix multiplier", *Applied Optics*, vol. 29, pp. 886 - 887, March 1990.

[3] C. J. Date, *An Introduction to Database Systems, Vol 1.*, Addison-Wesley, 1986

[4] Pericles A. Mitkas, P. Bruce Berra, "Implementation of Relational Database Operations in Optics," *in proceedings of the 25th Hawaii International Conference on System Sciences*, pp. 338 - 347, Hawaii, 1992.

[5] P. B. Berra, A. Ghafoor, P. A. Mitkas, S. J. Marcinkowski, and M. Guizani, "The Impact of Optics on Data and Knowledge Base Processing Systems," *IEEE Trans. on Knowledge and Data Eng.*, vol. 1, pp. 111 - 131, March 1989.

[6] K. W. Wong, L. M. Cheng, and M. C. Poon, "Design of Digital-Optical Processors by Using Both Intensity and Polarization-Encoding Schemes," *Applied Optics*, vol. 31, no. 17, pp. 3225 - 3232, June 1992

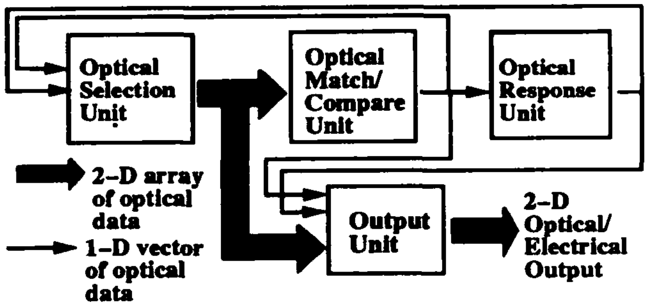


Figure 1. An overview of the Optical Content-Addressable Parallel Processor (OCAPP)

Legend:

- S/CL1 – Spherical/Cylindrical lens combination
- 2-D MCU – 2 Dimensional Match/ Compare Unit
- GE – Greater than or equals register
- LE – Less than or equals register
- L – Less than register
- G – Greater than register
- R – Equals register
- SLM – Spatial Light Modulator

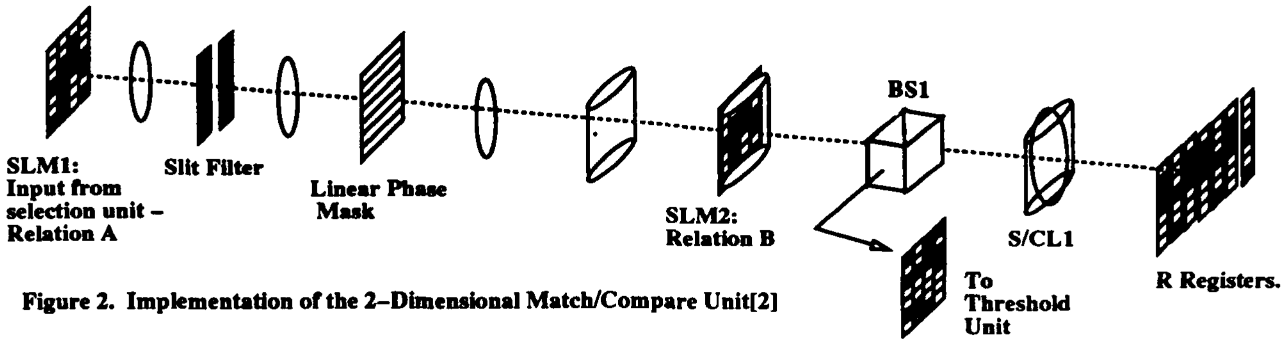


Figure 2. Implementation of the 2-Dimensional Match/Compare Unit[2]

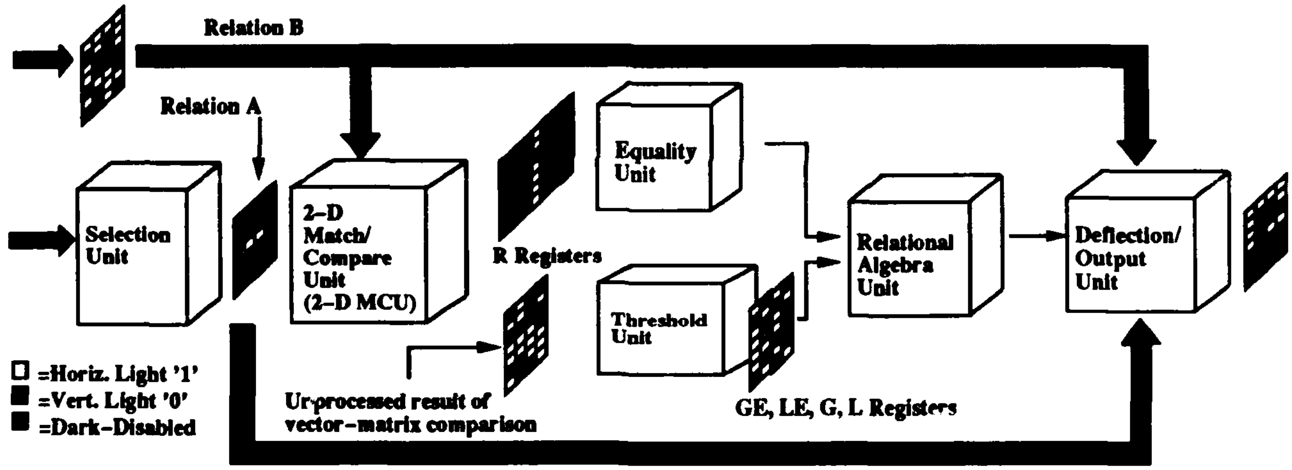


Figure 3. A schematic diagram of the OCAPP Database Machine

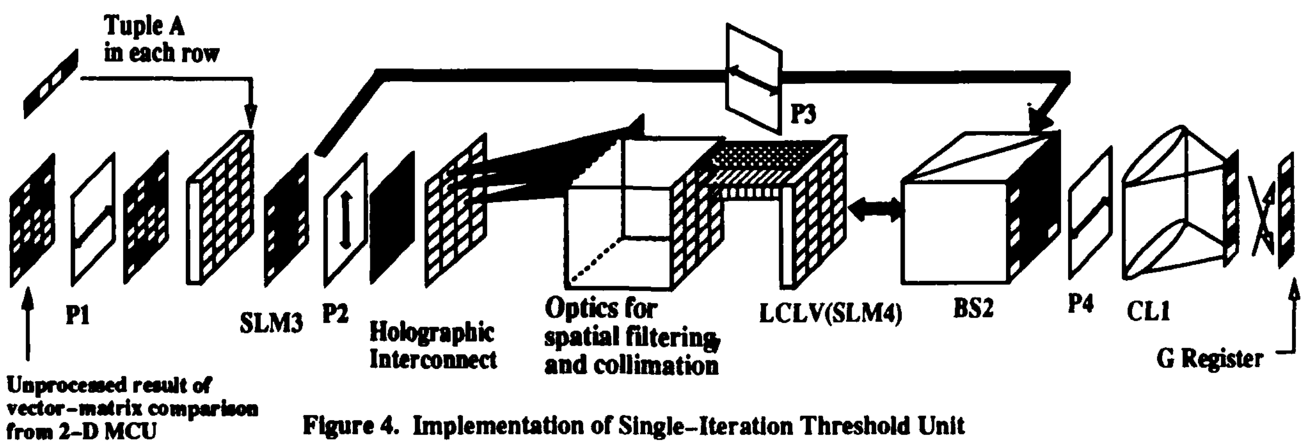


Figure 4. Implementation of Single-Iteration Threshold Unit

Volume Holographic Storage and Processing for Large Databases

Jeff Brown, L.J. Irakliotis, and Pericles A. Mitkas

Optoelectronic Computing Systems Center
and Department of Electrical Engineering
Colorado State University
Fort Collins, CO 80523
(303) 491-2021

1. Introduction

Today's most powerful electronic computers are capable of processing data at rates reaching or exceeding hundreds of GFLOPS. These rates can be achieved when the data involved reside in the main memory of the system which is often the case with numerical applications (matrix operations, etc). A different family of applications, however, that include data and knowledgebase management, rely on the computer's ability to process a vast amount of data in a real-time enviroment. For these input/output intensive applications the performance of the secondary memory system becomes critical and is usually determined by two factors: storage capacity and data transfer rates.

Among various secondary storage systems, holographic memories offer very large storage capacity, high-speed parallel data access, and low cost per bit. Access to the data is made by turning on selected pixels in a large array, which is an inertialess operation. In contrast to the serial output of magnetic and optical disks, the output of a volume hologram can be two dimensional, which combined with the absence of moving parts in the read/write mechanism, results in higher transfer rates that can reach hundreds of megabytes/sec. An additional advantage of holographic storage is its ability to operate in an associative addressing mode that is always preferable in a database environment. The above characteristics make holographic memories ideal candidates for database secondary storage [1, 2].

In this paper, we present the design of a volume holographic system for storage and processing of data in a relational database environment. Data records can be written and retrieved in parallel. The associative nature of holography in conjunction with a spatial data encoding scheme allows searches based on content, and thus, facilitates the implementation of several relational database operations. After the description of the system and its different access modes, the encoding scheme is discussed and the steps for the execution of selected relational operations are outlined.

2. Volume Holographic Storage for Relational Databases

The widely used relational database model was adopted because the tabular representation of data in a relation can be easily mapped on two-dimensional optical processing arrays or pages of a holographic memory. Figure 1 depicts a volume holographic database system (VHDS) that allows fully and partly associative searching as well as direct addressing. The database is stored in a Fe:$LiNbO_3$ crystal. Recent experiments have shown that, using angular multiplexing, 500-1000 holograms can be stored in a photorefractive medium with an acceptable bit-error rate [3,4,5]. Two spatial light modulators, SLM1 and SLM2, are required for recording new data and interrogating the database. Input patterns (data pages) are loaded into SLM1 and light from SLM2 provides the corresponding reference beams. SLM1 is two dimensional for the data, and SLM2 is one dimensional to emulate reference beams coming from different angles. Thus, data are stored using angular multiplexing. The output of the system is detected by two CCD arrays (CCD1 and CCD2). The electronic output of these arrays can be transfered to the host computer or fed back to the modulators for subsequent iterations. The entire process is supervised by a PC.

2.1 *Data Recording*

During the recording phase, a two-dimensional pattern, corresponding to a data page k, is loaded into SLM1. The k-th pixel is turned on at SLM2 to provide the reference beam associated with the pattern on SLM1 and is equivalent to the physical address of data page k. In the next step, a new pattern is loaded into SLM1, a different pixel is turned on at SLM2, and the new record is stored in the memory. The recording of two pages is illustrated in Figure 2. When the maximum number of holograms that can be superimposed in the same volume is reached, the crystal is rotated and data are recorded in a different part of the medium.

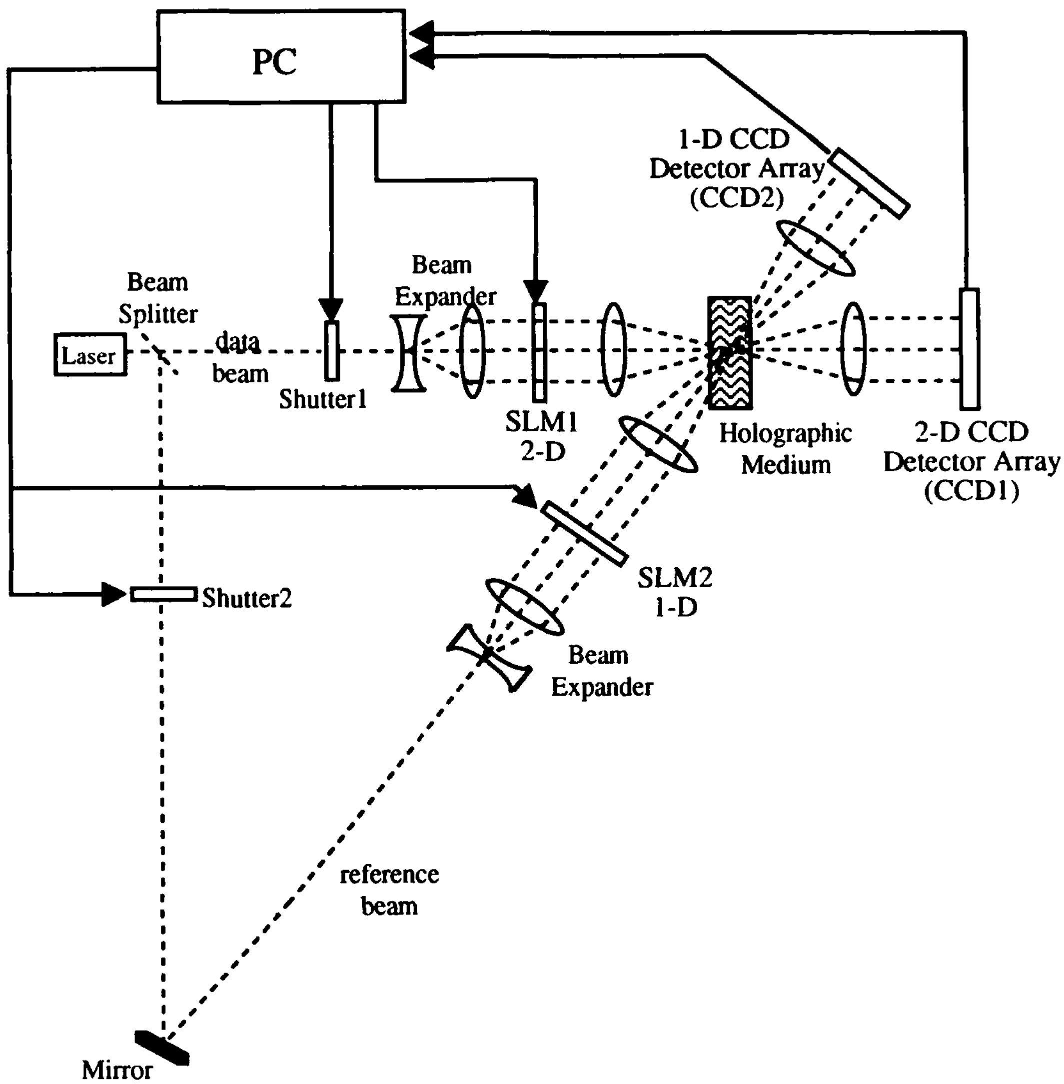


Figure 1. System diagram

2.2 *Encoding Scheme*

An efficient data encoding scheme must minimize the bit error rate and false responses during selective data retrieval. In a database environment, data are alphanumeric and, in most cases, a data item is either a letter string or a number. Furthermore, there is seldom a need to distinguish between upper and lower case letters. Therefore, the set of different character values can be restricted to about 50 elements (26 letters, 10 digits, and certain symbols and punctuation marks). If the length of a record is M characters (bytes), then a record can be represented as an array of Mx50 pixels. In each one of these arrays, only M pixels will be bright (on) at any given time. The row number *i* of a bright pixel will denote its alphanumeric value while the column *j* will indicate the position of the corresponding character in the record. A possible encoding scheme is illustrated in Figure 3, where all the zeros must appear in the first row, the ones in the second row, the A's in the 11th row, etc.

This encoding scheme introduces significant redundancy, but it is expected to minimize noise and false responses during the search and retrieval of data and allow for a certain degree of error detection. As an example, consider a scheme that restricts letters to odd rows only. Then, if a retrieved page contains a 1 in the pixel (3,*j*), where *j* is a column of a numerical (and not alphabetical) data field, we know that an error has occured.

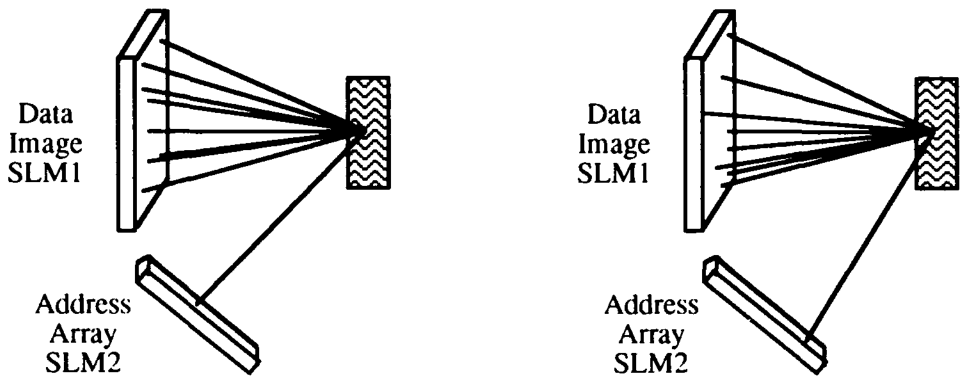


Figure 2. Recording of two pages on the hologram

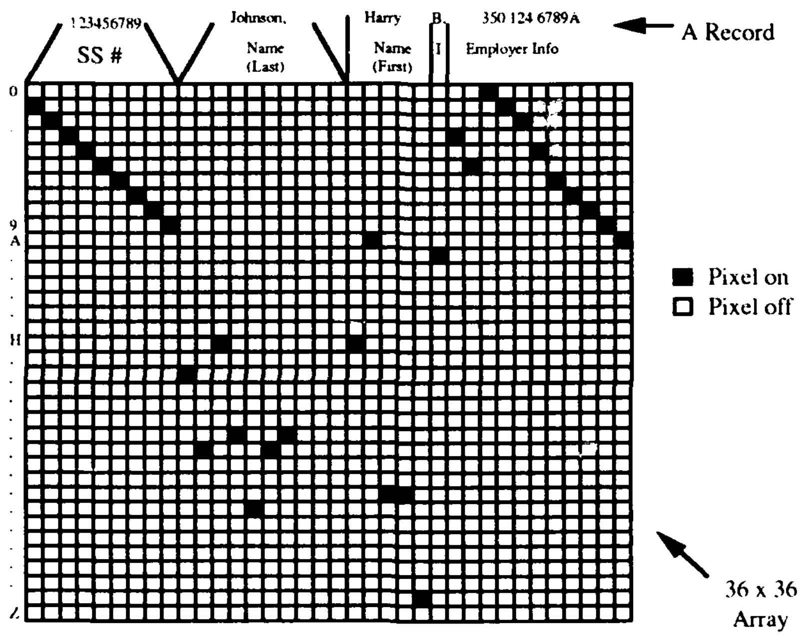


Figure 3. An example of data encoding

3. Data Retrieval

As previously mentioned, the system is able to retrieve data pages from the hologram based on physical address or content. Most of the database operations involve the comparison of a reference value to all the records in the database and can benefit by a content-addressable memory. On the other hand, the physical address is required when a specific record must be retrieved.

a) *Content-based mode*. During associative read, the search argument is loaded on SLM1 and illuminates the hologram. For every data page with an entry equal to the search argument, the pixel k of CCD2 will be bright, where k is the position of the reference source at SLM2 used for recording that particular data page. In a second step, the output of CCD2 can be fed back to SLM2 to retrieve the matching data pages.

b) *Address-based mode*. By turning on a pixel at SLM2, we can retrieve (on CCD1) the data page that was recorded with that particular reference beam. Only one data page can be retrieved in each step, but this process can be concluded very fast because it is bound by the response time of the photodetectors. Since SLM2 is one-dimensional, each one of its pixels can be addressed in paralell. As a result, the time to switch two pixels is significantly shorter than the time required to load an entire frame on a two-dimensional modulator. The addresses of the data pages will be provided by the output of CCD2 or the electronic control unit.

4. Relational Database Operations

A number of relational database operations can be executed directly with VHDS. The steps required for some of them are outlined below.

a) *Retrieval.* When all the data pages of a relation must be retrieved, the physical address of every page will be used and pages will be read out one by one, converted to electronic data, and decoded into ASCII characters.

b) *Projection.* This operation, which requires discarding of specific data fields (columns) of a data page, can be easily implemented during data retrieval by disabling the output of the corresponding CCD array cells. The system will operate in the address-based mode.

c) *Selection.* This operation will be concluded in two phases. First, the selection argument is loaded into SLM1, and an associative search is performed to identify the data pages that satisfy the selection criterion. Second, using the retrieved physical addresses, the qualifying data pages are read with the system operating in the address-based mode. A selection followed by a projection will not require any extra time because projection can be performed during retrieval.

d) *Binary operations(join, intersection, union, set difference).* These operations require input from two different relations that can be stored in two different areas of the hologram. The records of the first relation are retrieved and stored in the electronic buffer. Their contents will be used as the search arguments for an associative retrieval on the second relation. Depending on the operation, the data pages of the second relation may have to be retrieved sequentially during a third phase.

5. Conclusion

VHDS combines the large storage capacity and parallelism of volume holographic memories with smart data searching techniques and can provide an attractive solution to the problems associated with the management of very large databases. The critical parameters affecting the performance of VHDS are storage capacity, data rates, and overall response time to various queries. Assuming a page size of nxm, N superimposed holograms in the same area of the crystal, and S different areas, the total capacity is equal to nxmxNxS. For N=500, nxm=2^{16}, and S=100, the overall capacity can reach 3.25 Gbits. The data rate is a function of the frame rate of the photodetectors. If CCD1 has a frame period of T_D, then the burst transfer rate is (nxm)/T_D and for T_D = 1μsec, it can reach 65Gbits/sec. The response time of an operation can be expressed in terms of the number of address-based and/or associative data accesses required. For a selection operation (associative search followed by an address-based sequential retrieval), the response time depends on the selectivity factor of a relation which is equal to the percentage of records that satisfy the selection criterion. The speed up of this approach versus the exhaustive search is inversely proportional to the selectivity factor.

6. References

1. P.B. Berra, A. Ghafoor, P.A. Mitkas, S.J. Marcinkowski, and M. Guizani, "The Impact of Optics on Data and Knowledge Base Systems", IEEE Trans. on Knowledge and Data Engineering, Vol. 1, No. 1, March 1989, pp 111-132.

2. P.A. Mitkas and P.B. Berra, "Parallel Optical Memories for Very Large Databases," Proc. of the SPIE 37th International Symposium on Optical Applied Science and Engineering, Conference on Very Large Optical Memories, San Diego, July 1992, SPIE Vol. 1773c.

3. A. Chiou, invited paper on "Multiplexed Volume Holograms," at the SPIE 37th International Symposium, Conference on Very Large Optical Memories, San Diego, July 1992.

4. A. Lahrichi, K. Johnson, E. Maniloff, and G. Fredericks, "Signal-to-noise Limitations on the Number of Channels in Holographic Interconnection Networks," Journal of Optical Society of America, Vol. 9, No. 5, May 1992, pp 749-754.

5. F.H. Mok, M.C. Tackitt, and H.M. Stoll, "Storage of 500 High-resolution Holograms in a $LiNbO_3$ Crystal," Optics Letters, Vol. 16, No.8, April 15, 1991, pp 605-607.

Wavelength Domain Processing

OWA 8:30am-10:00am
Mesquite Room

Demetri Psaltis, *Presider*
California Institute of Technology

Spectral Hole-Burning, Holographic Storage and Processing

Urs P. Wild, Stefan Bernet, Stefan Altner, Eric S. Maniloff, and Alois Renn

Physical Chemistry Laboratory, Swiss Federal Institute of Technology, ETH-Zentrum
CH-8092 Zurich, Switzerland, Tel.: +41 1 256-4381; Fax: +41 1 252-3402

I. Introduction

The lowest electronic transitions of many absorbing centers in solid hosts consist of extremely narrow zero phonon lines distributed over a broad spectral range. Frequency selective modification - spectral hole burning (SHB) - of such inhomogeneously broadened bands has several potential applications. Using selective recording of holograms, SHB [1] allows for frequency multiplexed optical memories [2-4], and opens interesting prospects for high density optical information storage and processing. The principles of such a molecular processors have been described, demonstrating that multi-color recording materials can be used for frequency multiplexed image recording and simultaneously for parallel processing of the stored information [5].

II. Holographic image storage

Fig.1. Selected images of a video movie recorded in a chlorin doped polymer film.

Combining spectral hole-burning and holography [6,7] has led to frequency and electric field selective image storage [2-4] allowing the storage of thousands of images within a single polymer film. During recording, the beam of a tunable single mode dye laser is split into reference and object beams and the sample is exposed to the interference pattern of the two beams. The integrated diffraction efficiency and the transmitted signal are monitored by photomultipliers, and the images are recorded using a CCD camera. The addressing of the individual images is performed by adjusting the parameters "frequency" and "electric field" to the values used during recording. A video movie consisting of 2100 images was stored in a polymer film, corresponding to a sequence of 80 s duration. Four typical frames retrieved from individual holograms are shown in fig 1.

The electric field applied to the sample represents a further storage dimension and increased storage capacity [8].

III. Logical operations

For specific molecular systems a spectral hole can be observed to split when an electric field is applied to the sample, and the Stark components of closely spaced holes can be made to overlap [9]. For example, a pair of holograms can be stored in the frequency - electric field plane in an arrangement as shown in fig. 2, burnt at the same frequency at different electric field strengths. The burning coordinates where the stored images can be reconstructed separately are drawn as filled circles. The electric field dependent positions of the Stark components are indicated by dashed lines. The regions where the dashed lines intersect are marked by squares. At these positions a coherent superposition of the recorded images can be reconstructed, the result of which depends on the phase properties of the holograms. As shown in fig. 2 two images, a horizontal bar and a vertical bar have been recorded at different electric fields. The images can be reconstructed individually by adjusting the recording parameters, (ν, E_1) and (ν, E_2). The coherent superposition of the images can be accessed at the frequency positions, ν_1 or ν_2, and the electric field, $(E_1+E_2)/2$. The results are shown for a phase differences 0 and π. Constructive interference (phase 0) leads to an increase of the image intensity when the images overlap, the images are added. Destructive interference (phase difference π) results in a subtraction of the images. Logical operations corresponding to "AND" or "XOR" functions can be implemented when appropriate discrimination is applied.

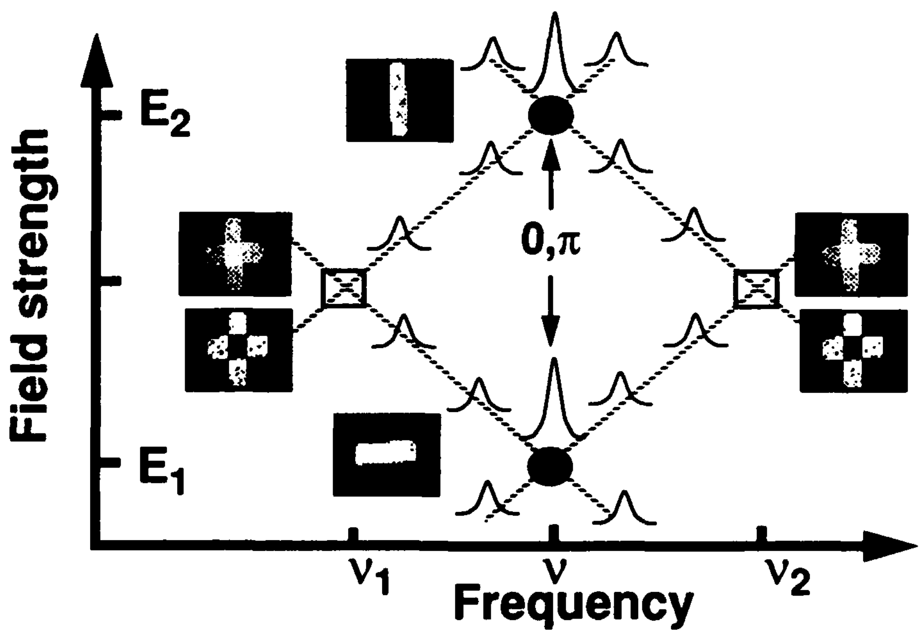


Fig 2. Logical operation with images. The recording positions are indicated by circles. Due to the Stark splitting of the spectral holes the holograms are made to overlap at the regions indicated by the squares. The reconstructed images show the result of the superposition: constructive interference, when a phase difference of zero was used and destructive interference when a phase difference of π was chosen.

IV. Molecular Computing

Logical operations corresponding to "AND" or "XOR" functions can be derived when appropriate discrimination is applied. In fig 3 a parallel addition using the AND as well as the XOR operations is shown. Each of the 8x8 patterns represents one bit of an array of 64 octal numbers, each encoded by a specific position within the grids. Bright areas symbolize a digital "1", dark areas correspond to "0". The patterns were transferred to the object beam by means of a liquid crystal display light modulator, and were stored in the hole burning material as holograms. The superposition of two images was performed in the material and the results have been obtained by appropriate thresholding subsequent to the read-out.

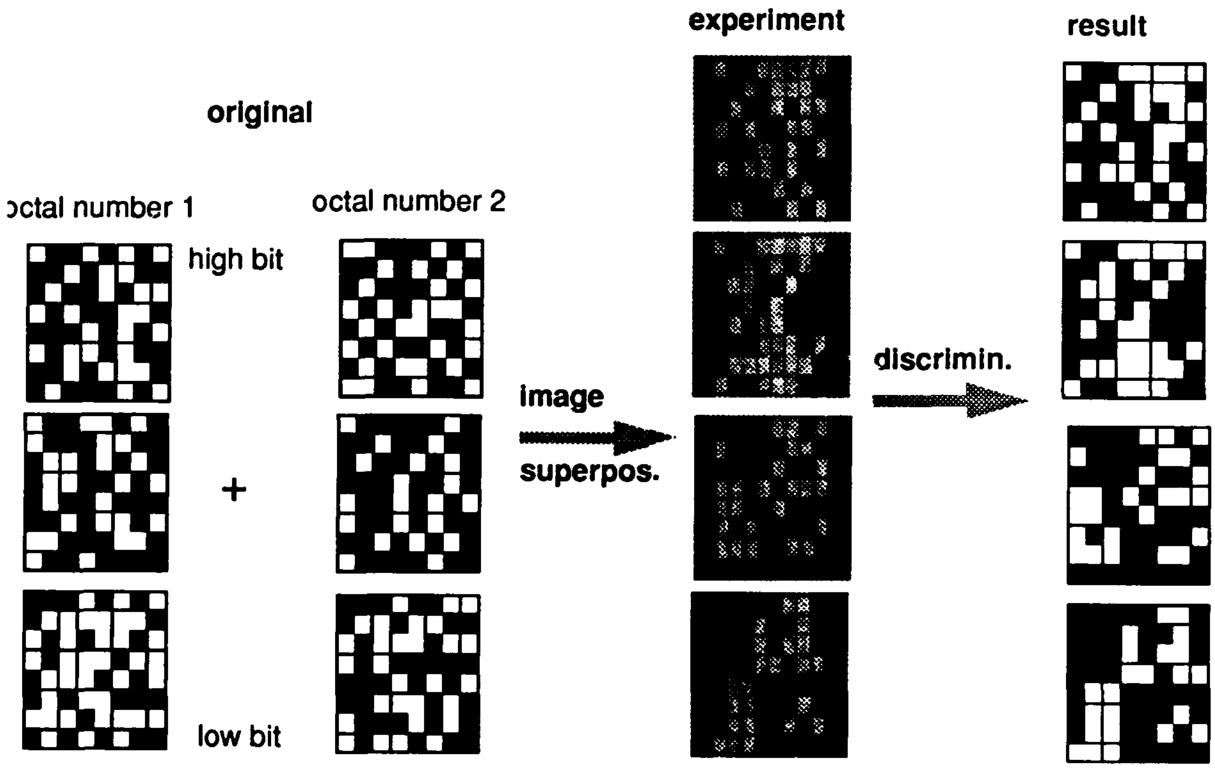


Fig 3.Parallel addition by means of spectral hole burning. *Two arrays of octal numbers have been stored and subsequently processed using a sequence of "AND" and "XOR" operations.*

V. Swept holograms

The diffraction efficiency of holograms considerably changes when different phase-shifts are applied as function either of the recording frequency [10] or the electric field applied to the sample. The hologram phase is controlled by using a piezo mounted mirror under computer control. Examples of this is shown in Fig. 4, which shows holograms swept in frequency and phase and E-field and phase. For frequency swept holograms the long range gratings are suppressed, which decreases crosstalk between the holograms. E-Field swept holograms produce asymmetric Stark-splittings in the field/frequency plane, with the direction of the splitting determined by the sign of the phase shift. In addition, E-field swept holograms can produce both constructive and destructive interference, depending on the field chosen for readout., which makes them ideal for coherent image superposition.

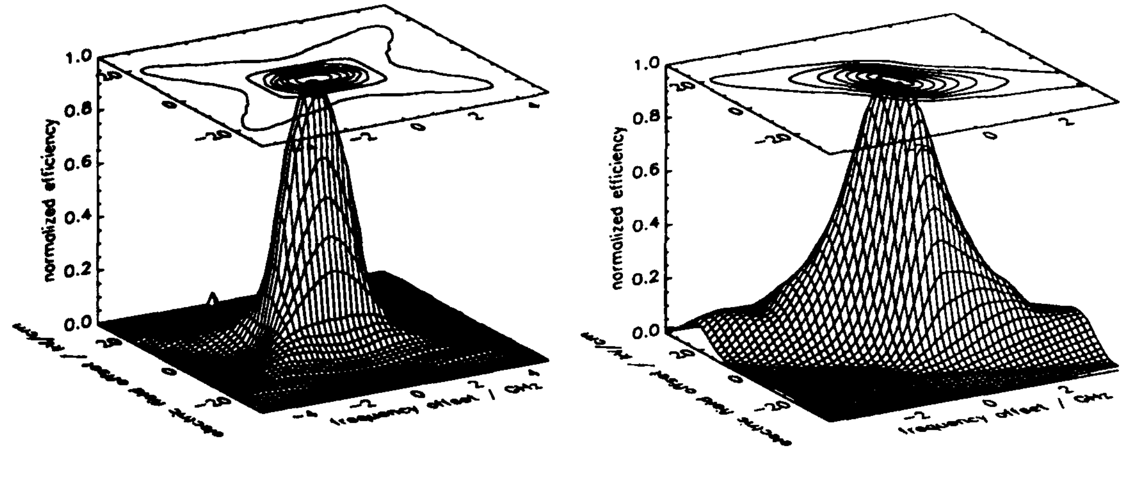


Fig 4. Surface and contour representation of a frequency - phase swept hologram (left) and an electric field - phase swept hologram (right). Note the different behavior in the wings of the signals.

VI. References

1 W. E. Moerner (Ed.), Persistent spectral hole burning: Science and applications, Topics in Current Physics **44**, Springer Verlag, Berlin, 1988.

2 U.P. Wild and A. Renn, "Molecular computing: A review I. Data and image storage" J. Mol. Electr. **7**, 1 (1991).

3 A. Renn and U.P.Wild, "Spectral hole burning and hologram storage", Appl. Opt. **26**, 4040 (1987).

4 C. De Caro, A. Renn, and U.P. Wild, "Hole burning Stark effect and data storage: 2. Holographic recording and detection of spectral holes. Appl. Opt. **30**, 2890 (1991).

5 U.P. Wild, A. Renn, C. De Caro, and S. Bernet, "Spectral hole burning and molecular computing" Appl. Opt. **29**, 4329 (1990).

6 F.A. Burkhalter, G. W. Suter, U. P. Wild, V. D. Samoilenko,N. V. Rasumova, and R. I. Personov, 'Hole Burning in the Absorption Spectrum of Chlorin in Polymer Films: Stark Effect and Temperature Dependence, 'Chem. Phys. Lett. **94** (1983) 483.

7 A. Renn, A.J. Meixner, U.P. Wild and F.A. Burkhalter, "Holographic Detection of Photochemical Holes," Chem. Phys. **93**, 157 (1985).

8 U.P. Wild, S. E. Bucher, and F.A. Burkhalter, "Hole Burning, Stark Effect, and Data Storage," Appl. Opt. **24** (1985) 1526.

9 A. Renn, A.J. Meixner, and U.P. Wild, "Hole burning and holography III. Electric field induced interference of holograms", J. Chem. Phys. **93**, 2299 (1990).

10 S. Bernet, B. Kohler, A. Rebane, A. Renn, and U.P. Wild, "Spectral hole-burning and holography. V. Asymmetric diffraction from thin holograms", J. Opt. Soc. Am. **B9**, 987 (1992).

Multiwavelength Logic Gates with High Computational Efficiency

D. J. Blumenthal and J. R. Sauer

Optoelectronic Computing Systems Research Center
University of Colorado at Boulder
Boulder, CO 80303
(303) 492-4858

1. Introduction

The wide bandwidth of optical communications channels offers the potential for extremely high throughput for interconnects in both computer communications and telecommunications applications. The utilization of this bandwidth typically involves conversion to electronic signals and processing using standard binary based techniques. We describe two classes of logic primitive which process information in the optical wavelength domain. It is shown that for implementation of truth table processors, these gates can achieve computational efficiencies many orders of magnitude greater than that obtained with Boolean logic. We also discuss the requirements and limitations associated with the described technique.

2. Multiwavelength Information Processing (MIP)

Multiwavelength information processing (MIP) is a technique whereby the complete temporal and optical bandwidth is available to communicate and process information. We focus on incoherent MIP where information is represented and processed based on the presence and absence of power at discrete optical wavelengths. There are two components to MIP: the representation and transfer of information, and the processing of information. The concept of MIP is illustrated through the design of two incoherent codes and compatible lookup table processors. The codes, described below, are illustrated by example in Fig. (1). $(U_\lambda; B_t)$: Unary in wavelength and binary in time. Any combination of d time slots and only one wavelength per time slot. $(B_\lambda; B_t)$: Binary in both wavelength and time. Any combination of d time slots and any combination of R wavelengths per time slot.

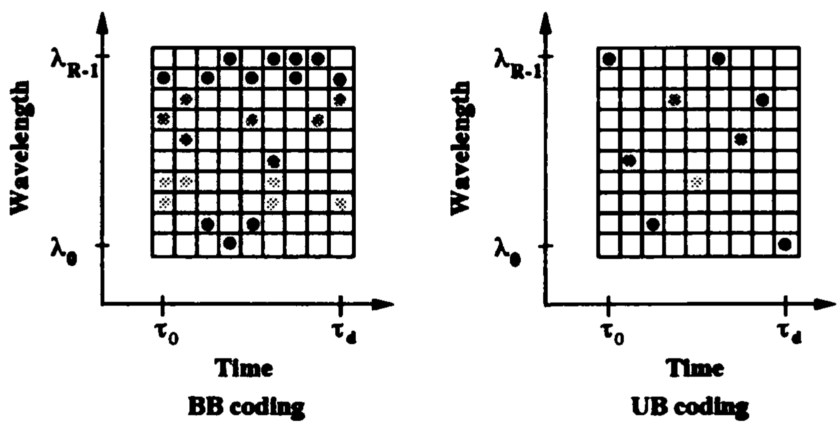


Figure 1: Examples of UU and BB codes in the time and wavelength domains

Quantities of interest are the information capacity and gain. Assuming that the information channel is noiseless, the information gain is defined as the ratio between the information capacity for and MIP code and a temporal binary code. Table 1 summarizes the information capacity, degrees of freedom, and information gain for the two codes. One class of number representation which benefits from MIP is multivalued logic (MVL). MVL systems represent and process information in a number radix system ≥ 2. If the optical equivalent of a multilevel system uses multiple power levels to represent a higher radix number, the minimum spacing between levels will be defined by shot and thermal noise, resulting in a noise/speed tradeoff. If instead we use discrete optical wavelengths to represent the multiple values, the limitations are based instead on phase noise and wavelength stability of the source.

Code	Degrees of Freedom	Information Capacity	Information Gain
UB	R^d	$d \log_2 R$	$log_2 R$
BB	2^{Rd}	Rd	R

Table 1: Summary of coding information results

3. Optical Set Logic (OSL)

Optical set logic (OSL) processes multiwavelength data as binary coded multi-valued numbers. We consider numbers which are represented using R wavelengths in a single time slot ($d = 1$). Multiple single-slot inputs to a processor are equivalent to $d \geq 1$. The representation of 2^{dR}-radix numbers using binary encoded electronic frequencies has been reported [YAH91]. After completion of the work reported in this paper, logic gates which use multiple optical wavelengths to code set logic functions different than those described here were reported [MAH92].

A set of R wavelengths $\{\Lambda\}$ represents a 2^R-radix number. An input variable is represented by $\{X_i\} \in \{\Lambda\}$. The complete absence of optical wavelengths is given by the null set $\{\phi\}$. We define three optical set logic primitives following [YAH91]: the OSD primitive is the set difference operator ($-$) in Eqn. (1) and effectively filters a fixed set of wavelengths given by a local constant $\{A\}$ from an arbitrary input $\{X_i\}$. The OSG in Eqn. (2) is simply an inverted switching function which passes a set of local wavelengths if there are no wavelengths present at the input, and has no optical output when any wavelengths are present at the input. The OSU gate in Eqn. (3) is realized using simple incoherent optical summation. In order to make cascadable gates, we also desire level restoration which can be implemented using a broadband saturating optical amplifier. These logic primitives map well to optical and photonic implementation.

$$
\begin{aligned}
OSD(\{X_i\};\{A\}) &= \{X_i\} - \{A\} = \{X_i\} \cap \overline{\{A\}} && (1)\\
OSG(\{A\};\{X_i\}) &= \begin{cases} \{A\} & \text{if } \{X_i\} = \{\phi\} \\ \{\phi\} & \text{otherwise} \end{cases} && (2)\\
OSU(\{X_1\},\{X_2\},\ldots,\{X_n\}) &= \{X_1\} \cup \{X_2\} \cup \ldots \cup \{X_n\} && (3)\\
\overline{\{X_i\}} &= OSD(\{\Lambda\};\{X_i\}) = \{\Lambda\} - \{X_i\} && (4)
\end{aligned}
$$

A technique for computing the complement of an optical set logic variable, which requires significantly less gates than that reported in [YAH91] is described. A dynamic OSD gate, where the local value is the input and the input is replaced by an optical source with all wavelengths in $\{\Lambda\}$. The complement is therefore given by the set difference function as in Eqn. (4). This type of gate can be realized with a device such as the acoust-optic tunable filter [Smi92].

An example lookup table is shown in Fig. (2). The constants A_{ij}, B_{ij}, and k_i are determined by a logic minimization technique described in [AG68]. The output of this processor has 2^R degrees of freedom. The degrees of freedom at the output can be increased by using multiple lookup tables to generate a multiple variable 2^R-radix output.

4. λ-gate logic

The special case of a unary code in a single time slot reduces to wavelength position modulation (WPM). WPM can be treated as an R-valued number and processed using a truth table processor constructed with the optical equivalent of a T-gate [Hur84]. We call these λ-gates, and as with T-gates they multiplex local constants according to the value of a data select, effectively routing the correct truth table values to the output. For example, a ternary λ-gate transfer function is written as $\Lambda(\lambda_D : \lambda_A, \lambda_B, \lambda_C)$, implying that the output is a function of the wavelength data select λ_D given a set of local constants λ_A, λ_B, and λ_C. For an R-radix system, the λ-gate is operationally an Rx1 multiplexer. The data select and local constants can

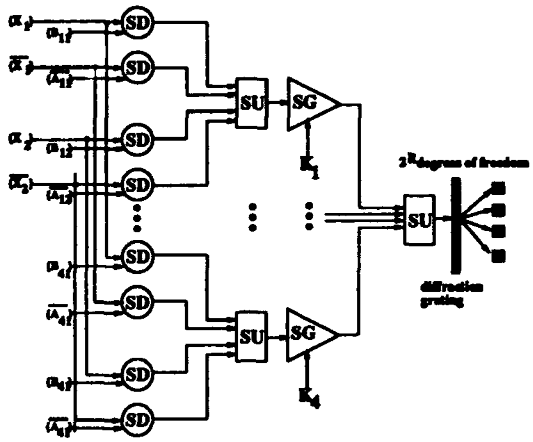
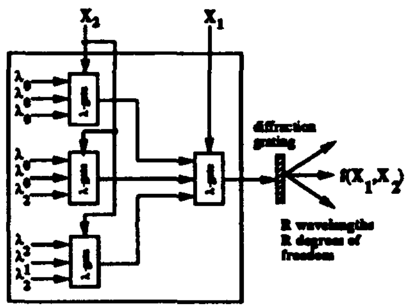


Figure 2: A two input routing table using the OSD, OSG, and OSU logic primitives

take on values of any one of the R wavelengths $\{0, \lambda_1, \ldots, \lambda_{R-1}\}$. The truth table and λ-gate realization for a ternary two input processor $f(X_1, X_2)$ is shown in Fig. (3).



X_1	X_2	$f(X_1,X_2)$
λ_0	λ_0	λ_0
λ_0	λ_1	λ_0
λ_0	λ_2	λ_0
λ_1	λ_0	λ_0
λ_1	λ_1	λ_0
λ_1	λ_2	λ_2
λ_2	λ_0	λ_2
λ_2	λ_1	λ_1
λ_2	λ_2	λ_2

Figure 3: An example λ-gate lookup table processor and truth table

5. Processing Efficiency

We define the gate efficiency in terms of the number of possible mappings in a truth table performed per gate. In this paper we assume that the cost or complexity of each gate is the same. While not a complete analysis, this simplification gives us an idea of how the efficiency of the architecture scales with increasing radix. The number of gates necessary to realize an OSL truth table processor, assuming that dual rail logic is used, is $\Gamma = (d+1)2^{d+1} + 1$. Similarly, the total number of λ-gates required are $\Gamma = \frac{R^d - 1}{R-1}$. The efficiency of an *OSL* based processor is given by Eqn. (5) and the efficiency of a λ-gate processor is given by Eqn. (6).

$$\varepsilon_{osl} = \frac{(2^R)^{2^{R^d}}}{(d+1)2^{d+1} + 1} \tag{5}$$

$$\varepsilon_{\lambda-gate} = \frac{(R-1)(R)^{R^d}}{R^d - 1} \tag{6}$$

The optical set logic and λ-gate efficiency in terms of number of mappings per gate, are plotted in Fig. (4) as a function of the number of wavelengths, or radix, R for single and double variables or time slots. For Boolean gates, the computational efficiency is approximately unity. The extremely high computational efficiency must be accompanied with several caveats. First, these numbers represent arbitrary truth tables where no logic minimization has occurred. Secondly, we assume that all degrees of freedom are usable at the

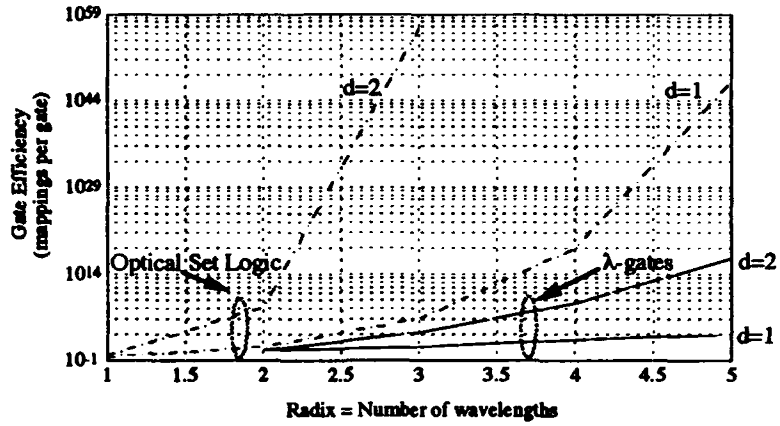


Figure 4: Gate computational efficiency curves for various combinations of R and d

output for some computational or control process. Third, the complexity of an MIP gate should be much less than if its function were realized using Boolean gates.

6. Summary

In this paper we describe the concept of multiwavelength information processing (MIP) and two MIP coding techniques. It is expected that this type of processing can result in improved computational efficiency and parallelism over that achievable with electronics. We illustrate this concept through the design of two lookup table processors based on *Optical Set Logic* and $\lambda - gate$ primitives. The resulting processors exhibit an extremely high computational efficiency in terms of number of mappings per gate. They also exhibit characteristics which are amenable to optical and photonic implementation. A primary goal of this work is to investigate high functionality gates which exploit the parallelism and wide bandwidth of multiwavelength systems. In the next stage of this work, we plan to compare MIP gates to the computational equivalent of electronic Boolean gates according to some measure such as volume and power dissipation. This increase in functionality has the potential to compensate for the larger footprint of optical devices over their electronic counterparts. Ultimately, the success of any technique will rest on a collection of issues: information capacity, performance, complexity, realizability, scalability, energy and computational efficiency, and cost.

7. Acknowledgements

This work is funded by the Air Force Office of Scientific Research under grant number AFOSR-91-0392.

8. References

[AG68] C. M. Allen and D. D. Givone. A minimization technique for multiple-valued logic systems. *IEEE Trans. on Comput.*, C-17:182–184, February 1968.

[Hur84] Stanley S. Hurst. Multiple-valued logic — its status and its future. *IEEE Transaction on Computers*, c-33(12):1160–1179, December 1984.

[MAH92] S. Maeda, T. Aoki, and T. Higuchi. Set-valued networks based on optical wavelength multiplexing. In *Proceedings of the IEEE 22st International Symposium on Multiple-Valued Logic*, pages 282–290, June 1992.

[Smi92] D.A. Smith. Tunable optical filters for lightwave networks. pages 19–20, Santa Barbara, CA, Aug 1992. IEEE/LEOS Topical Meeting on Optical Multi-Access Networks.

[YAH91] Y. Yuminaka, T. Aoki, and T. Higuchi. Design of a set logic network based on frequency multiplexing and its applications to image processing. In *Proceedings of the IEEE 21st International Symposium on Multiple-Valued Logic*, pages 8–15, August 1991.

Multiwavelength Optical Half Adder

Pochi Yeh, Scott Campbell and Shaomin Zhou

Electrical and Computer Engineering
University of California
Santa Barbara, CA 93106
805-893-3981, 805-893-7015

In the past decade there have been many proposals for and demonstrations of digital optical computing architectures [1-7]. Some of these systems offer great potential in terms of computing speed. However, none of them have utilized the great benefit optics offers in terms of bandwidth, even though it has been proposed that multiple wavelengths can be employed to transmit multivalued information in an optical fiber [8]. Optical computing systems that utilize multiple wavelengths as an added dimension to code information will enjoy greatly increased capacity and data transmission speed.

To capitalize on the advantages of optical bandwidth, multiwavelength information processing (MIP) was proposed to code, transfer, and process information [9]. For example, in MIP an 8-bit word can be represented by the presence or absence of optical power at any of 8 discrete wavelengths, each representing one bit in the word. To achieve parallel computation, nonlinear optical logic gates that perform independent logic operations on individual wavelengths are needed. In this paper, we propose and demonstrate several unique concepts which utilize four-wave mixing in photorefractive media for the implementation of a multiwavelength half adder. This half adder is accommodated through iterations of CARRY and SUM operations.

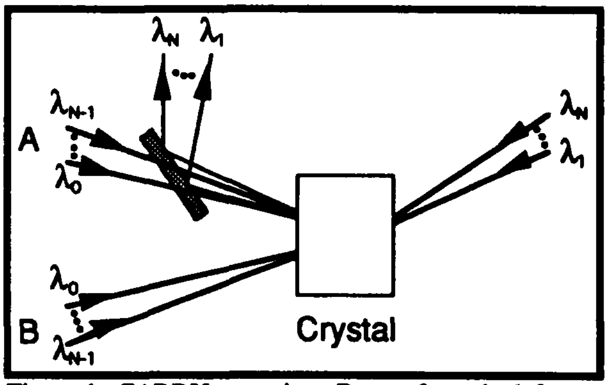


Figure 1. CARRY operation. Beams from the left at λ_{M-1} write index gratings that are read by beams from the right at λ_M. The output exits the system upwards.

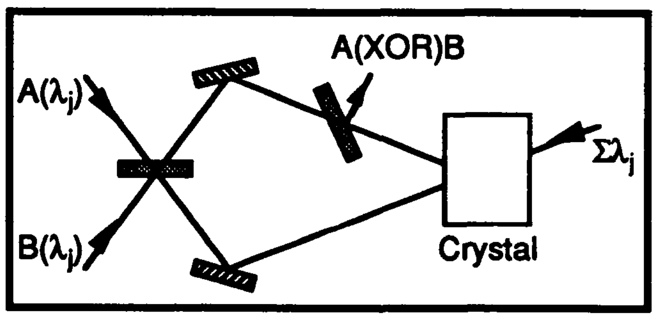


Figure 2. SUM operation. Beams from port A write index gratings that are π shifted from those written by beams from port B. The net gratings are read by an all wavelength beam set from the right. The resulting diffraction exits the system upwards.

In our system, multiwavelength phase matched four-wave mixing is employed to achieve the spectrally parallel operations of CARRY, and multiwavelength holographic interference is employed to achieve the spectrally parallel operations of SUM. Since a CARRY operation is equivalent to an AND gate with a one-bit shift in registration, we utilized simple two-beam interference to write gratings in a photorefractive media with beams of one wavelength (inputting

A(AND)B), and then tested for the existence of such gratings by reading them out with another wavelength (to allow the registration shift), as shown in Figure 1. Here, the two input write beams would be at λ_{M-1} and the read beam would be at λ_M. To implement the SUM operation, a Mach-Zehnder interferometer (MZI) was used in conjunction with a photorefractive crystal to simulate an XOR gate, as shown in Figure 2. The inputs to this gate were again the beams A and B, each entering one port of the interferometer. As a result of time reversal symmetry in the MZI, a π phase shift exists between the gratings written by beams from port A and those written by beams from port B [10-11]. Thus, a net grating exists only if a beam from one port or the other is present, while no grating exists in the absence of beams or in the simultaneous presence of beams from both ports at the same wavelength. Here, the two input write beams and the read beam would all be at λ_M. This produces the XOR gate necessary for the SUM operation.

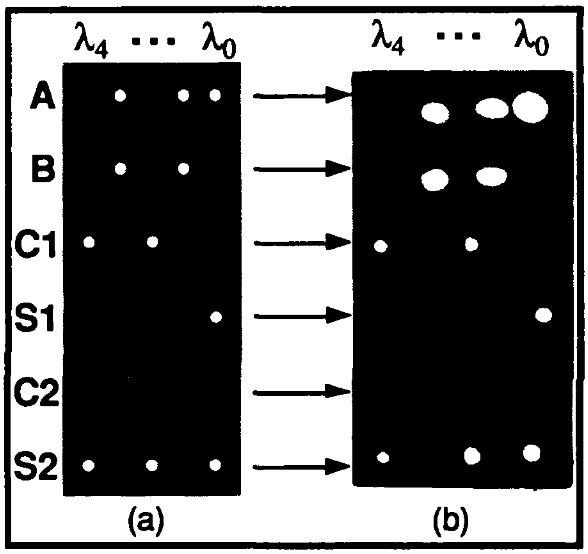


Figure 3. (a) shows a proposed addition operation for inputs A and B, along with the corresponding CARRY and SUM steps leading to the solution. (b) shows the corresponding experimental results.

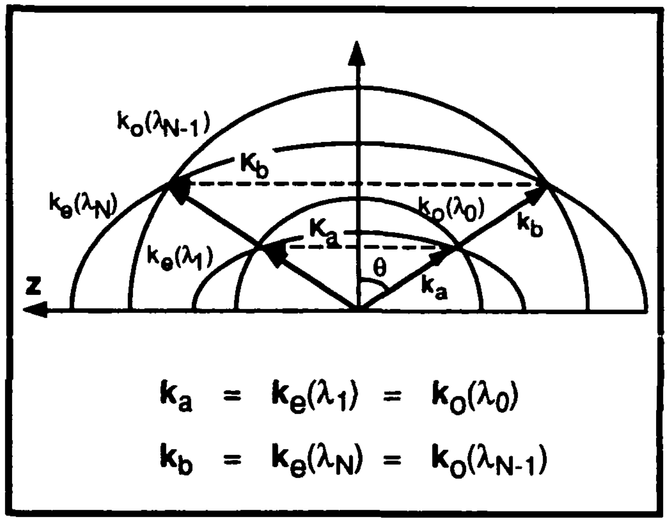


Figure 4. Anisotropic collinear diffraction. Here, gratings written by ordinary light at λ_{N-1} are read out by extra-ordinary light at λ_N. All wavelengths in a given beam set enter collinearly. Collinear readout then also results.

To describe the principle of operation of the multiwavelength optical half adder, we consider the problem of adding two binary coded numbers, A=(01011) and B=(01010), as given in Figure 3a. A 5-bit binary coded word is represented with beams from 5 discrete wavelengths, λ_4, λ_3, λ_2, λ_1, λ_0, arranged from the most significant bit to the least significant bit. In the first step of the addition process, an AND operation on A and B with a shift in registration leads to a CARRY output C1=(10100), while an XOR operation on A and B leads to a SUM output S1=(00001). In the second step, an AND operation on C1 and S1 with a shift in registration leads to a CARRY output C2=(00000), while an XOR operation on C1 and S1 leads to a SUM output S2=(10101). Further iterations of this process will yield identical results, so S2=(10101) is taken as the final answer. It is seen that a wavelength conversion is needed to perform the CARRY operation. That is, if λ_3 and λ_1 are present in both inputs A and B, then λ_4 and λ_2 will be present in the CARRY output (the λ_3 and λ_1 input beams are thus converted to output beams at λ_4 and λ_2, respectively). Nonlinear optical devices are ideally suited for such wavelength (or frequency) conversion via Bragg matched diffraction, where a beam at λ_N can read a grating written by beams at λ_M (M = N-1 in the CARRY, and M = N in the SUM). One must remain cautious, though, to maintain spectral independence so as to avoid readout cross talk between bits of different wavelength (e.g., in the CARRY operation, a beam at λ_N should only read gratings written by beams at λ_{N-1}). Optical four-wave mixing in photorefractive media ideally fulfills these requirements of wavelength conversion and spectral independence.

In multiwavelength four-wave mixing in a photorefractive media, the input beam sets $A(\lambda_4, \lambda_3, \lambda_2, \lambda_1, \lambda_0)$ or $B(\lambda_4, \lambda_3, \lambda_2, \lambda_1, \lambda_0)$ can enter the crystal in a number of ways. One way would be to allow the write beams at λ_M to each enter the crystal at unique angles θ_M. The readout beams at λ_N would then enter the crystal at the appropriate phase matching angles θ_N, as in Figure 1. Another way would be to have each input beam set propagating collinearly, as they would if they had been transported down an optical fiber. In this manner, even though the various wavelengths from a given set have the same crystal entry angle, the respective index grating wave numbers written between the beams from A and B would all still be unique since each of the $(\lambda_4, \lambda_3, \lambda_2, \lambda_1, \lambda_0)$ are unique. Performing the readout process for the SUM operation (where the input beam sets A and B enter the two input ports of a Mach-Zehnder interferometer) would then be a matter of simply counter propagating the beams from one arm of the interferometer with a beam containing all of the $(\lambda_4, \lambda_3, \lambda_2, \lambda_1, \lambda_0)$ also propagating collinearly. In the CARRY operation, one could read index gratings in one of two ways. The first way would be to angularly tune the readout beam set wavelengths at λ_N to phase match the index gratings written by the write beams at λ_{N-1}. In this manner, the output beam set would not be collinear, which would be inconvenient for coupling into waveguides. A second approach would be to utilize orthogonally polarized read beams in an optically anisotropic media.

Referring to Figure 4, we consider the momentum matching (Bragg) condition in K-space for a negative uniaxial crystal with ordinary write beams and extraordinary read beams. In this arrangement, the wavevector circle for write beams at λ_{N-1} must intersect the wavevector ellipse for a read beam at λ_N where phase-matched readout is to occur from a grating wavevector K_{N-1}. From a given angle of incidence θ, and a knowledge of the crystal's refractive index variations as a function of wavelength and polarization state, the entire set of write and read wavelengths are determined once one particular wavelength is chosen. For example, in SBN at an internal angle of $\theta = 70$ degrees, choosing $\lambda_0 = 500$ nm forces $\lambda_1 = 499.67$ nm, which forces $\lambda_2 = 499.34$ nm, which forces $\lambda_3 = 499.01$ nm, which forces $\lambda_4 = 498.68$ nm. These wavelengths are approximately 0.33 nm apart. For practical applications, the total wavelength spread, $\lambda_{LSB} - \lambda_{MSB}$, must be small enough so as to avoid significant propagation phase delays due to group velocity dispersion if fibers are to be used, yet large enough to avoid significant crosstalk between various readout wavelengths. To satisfy these restrictions, we suggest the use of reflection gratings in a photorefractive media such as SBN.

To demonstrate the CARRY and SUM operations, we utilized the experimental set ups shown in Figures 1 and 2, respectively. Our light source was an Argon ion laser operating multiline. From this we chose the five strongest lines such that $\lambda_0 = 514.5$ nm, $\lambda_1 = 496.5$ nm,

$\lambda_2 = 488.0$ nm, $\lambda_3 = 476.5$ nm, and $\lambda_4 = 457.9$ nm. All of these beams were extraordinarily polarized. The average coherence length for any given line was about 12 centimeters. SF10 prisms were utilized to separate the wavelengths, and shutters were used to determine a given input beam's logic state. Our photorefractive crystal was 4% MgO doped lithium niobate, about 8 millimeters on each side. In both the CARRY and the SUM operations, the external full beam write angle was 25 degrees. For the CARRY operation, neutral density filters were used to balance the powers of all of the read and write beams to about 40 mW each. An incoherent optical bias was introduced to the crystal so as to reduce the output signal strength variations as various write beams were turned on or off. Resulting read beam diffraction efficiencies were around 1% for index grating modulation depths around 4%. To implement the SUM operation, we utilized a Mach-Zehnder interferometer (MZI) XOR gate as shown in Figure 2. Specific wavelengths chosen to be present in the input beam sets $A(\lambda_4, \lambda_3, \lambda_2, \lambda_1, \lambda_0)$ or $B(\lambda_4, \lambda_3, \lambda_2, \lambda_1, \lambda_0)$ were selected using prisms and shutters, while collinear propagation of the beams in a given beam set was achieved by retro-reflecting each beam after it passed through a prism and a shutter. The net grating set in the XOR gate was read by a beam containing all wavelengths propagating collinearly, with a total power of 500 mW, while the total write beam power was 100 mW. The distributed power between the wavelengths was that which exited the laser. Diffraction efficiencies were again around 1% when a beam from one port was on, but about 0.04% if beams from both ports were on. Figure 3b shows a multiple exposure photograph of the various stages of half adder operation in accordance with the states specified in Figure 3a.

In conclusion, we have proposed, analyzed, and demonstrated a mulitwavelength optical half adder. To our knowledge, it is the first demonstration of digital optical computing utilizing spectral parallelism in the optical domain. It demonstrates the use of photorefractive materials to achieve CARRY and SUM operations via optical four wave mixing.

References:

1. See, for example, Special issue on Optical Computing, Proc. IEEE, 72 (1984).
2. A. W. Lohmann and J. Weigelt, "Digital optical adder based on spatial filtering," Appl. Opt., 25, 3047 (1986).
3. R. A. Athale and S. H. Lee, "Development of an Optical Parallel Logic Device and a Half-Adder Circuit for Digital Optical Processing," Opt. Eng., 18, 513 (1979).
4. J. Tanida and Y. Ichioka, "Optical Logic Array Processor using Shadowgrams, "J. Opt. Soc. Am., 73, 800 (1983).
5. B. S. W. Herrett, "All Optical Computation - a Parallel Integrator Based upon a Single Gate Full Adder, "Opt. Commun., 56, 87 (1985).
6. M. M. Mirsalehi, C. C. Guest, and T. K. Gaylord, "Residue Number System Holographic Truth-Table Look-Up Processing: Detector Threshold Setting, and Probability of Error due to Amplitude and Phase Variations," Appl. Opt., 22, 3583 (1983).
7. M. T. Fatehi, K. C. Wasmundt, and S. A. Collins, Jr., "Optical Logic Gates using Liquid Crystal Light Valve: Implementation and Application Example," Appl. Opt., 20, 2250 (1981).
8. S.L. Hurst, "Fiberoptics, a multiple-valued interconnection means ?" IEEE 10th International Symposium on Multiple-Valued Logic, pages 115-119, June (1980).
9. See, for example, D. J. Blumenthal and J. R. Sauer, "Multiwavelength information processing in Gigabit photonic switching networks," SPIE Proc. Vol. 1787, paper 04 (1992).
10. P. Yeh, T.Y. Chang and P.H. Beckwith, "Real-Time Optical Image Subtraction using Dynamic Holographic Interference in Photorefractive Media," Opt. Lett. 13. 586-588 (1988).
11. P. Yeh, D. Zhang and C. Gu, "Parallel Subtraction of Fourier Power Spectrum using Holographic Interferometry," Optics Lett. 17, 70-72 (1992).

Wavelength Multiplexed Computer Generated Volume Holography

Joseph Rosen,

Rome Laboratories, Hanscom Airforce Base, MA 01731,

Mordechai Segev and **Amnon Yariv,**

California Institute of Technology, 128-95, Pasadena, CA 91125

We demonstrate the first experiment of multiple-storage of computer generated holograms (CGHs) in a volume holographic storage medium. The motivations for this direction are (I) realization of an interface between an electronic computer and a volume holographic data storage, and (II) storage of artificial data, which does not exist as a two or three dimensional object (for example, phase-only filters). In the first application the use of CGHs gives a useful degree of freedom that introduces adaptivity in the holographic-recording process, and enables a user to compensate for imperfections in the recording system (while keeping the desired data unchanged). The second provides a convenient realization of complex filters for image-processing purposes. The field of computer generated volume holography was pioneered by Pugliese and Morris(1) who stored a single CGH in a photorefractive (PR) crystal. Their results suffered from problems associated with thin holograms(2), such as high diffraction orders, which were eliminated only at the expense of resolution in the reconstructed data page.

We demonstrate the loading of planar holographic data from a computer Random Access Memory (RAM) into a holographic storage medium (a PR crystal), and subsequent conversion to volume holograms which display all the advantages of volume holography(2). Furthermore, we suggest a method of to speeding-up the data loading process, by encoding more than one hologram at a time on the "data converter" (electronic to optical), i.e.

on the SLM, and transforming it to multiple volume holograms, recorded simultaneously in the volume storage medium. Finally, we demonstrate a convenient method for conversion of wavelength multiplexed holograms into angular multiplexing.

Our experimental setup is sketched in figure 1. The data-loading process starts in the computer RAM. A data page in RAM is defined as a digital matrix or, alternatively, by an analog image. The data is transformed into a hologram, designed to be borne on a known "carrier" light beam, and to be reconstructed by another known readout beam. The SLM used in this experiment is a magneto-optic SLM with 128x128 pixels. Since it is a binary low resolution SLM, we find that coding a CGH by the Projection Onto Convex Set (POCS) algorithm(3) is a rapid and efficient method. The CGH is transmitted from the computer to the SLM, and the light beam illuminating it reconstructs the data. Since the SLM is a thin hologram, the light diffracts off it into multiple orders of diffraction, each (apart from the zero order) bears all the data.

An essential issue is the conversion of the computer generated planar hologram into a volume hologram. Incoherent imaging of the hologram from the SLM onto the volume recording medium, may change the distribution of the information throughout the volume (owing to a varying fringe visibility along z), but the overall contents of both holograms are in principle identical. The immediate implication is high diffraction efficiency at the expense of resolution or vice-versa(1).

We demonstrate a method for converting a planar hologram into a volume hologram, by utilizing the diffraction properties of a coherent image-bearing optical beam. The principle is that all the data is borne on each one of the diffraction orders of the reconstructed planar hologram from the SLM. Therefore, imaging the CGH from the SLM onto the storage medium is not necessary, and one may use a single diffraction order with no loss of information. The zero order of diffraction reacts as the reference in the recording process, and the resultant interference grating is tilted with respect to the optical axis. The stringent condition for imaging the CGH onto the storage medium is therefore relaxed to an overlap requirement of the zero and, say the first, diffraction orders. Good visibility is hence displayed over a range that is much larger than the focal depth of the imaging system.

We pursue the properties of volume holography, by demonstrating simultaneous recording and reconstruction of two wavelength-multiplexed holograms. The original "data pages" were the letters S and T, and their computer-generated holograms were carefully designed, in accordance with Bragg and Snel laws, to be reconstructed each by a different wavelength: (T by a red HeNe laser, and S by an Ar laser at $\lambda=0.514\mu m$), from the **same angular direction** (see Figs. 1,2a). The holograms were superimposed, transmitted to the SLM, and imaged onto the PR crystal ($BaTiO_3$). The imaging system de-magnified the CGH to a density of 1280 lines/mm in the PR crystal, yielding a wavelength selectivity of $\Delta\lambda \cong 1Å$ for interaction length of $L \cong 7mm$. The reconstruction was performed by the appropriate laser beams, which were incident upon the crystal from the same direction. Since both holograms were converted into the "tilted" volume holograms simultaneously during the recording process, each one of them acquires a different tilt angle, due to the different periodicity. The reconstructed images emerge *with an angular separation* of 1.4^o, owing to the combined effect of a different tilt angle and Bragg matching in the presence of chromatic dispersion (n varies with λ). Fig. 2a shows the diffraction from the tilted volume gratings, and demonstrates the dual wavelength and angular separation.

The experimental results are shown on Figs. 2b-d. Figure 2b shows the "double" computer-generated hologram, and Figs. 2c shows the direct reconstruction from the SLM with the HeNe laser only. The SLM acts as a "sampler" due to its binary grid, and by spatial filtering we stored **four** holograms simultaneously, each yielded one letter of the four located on the right hand of the optical axis (see Fig. 2c). Being a thin hologram, the SLM does not possess the wavelength selectivity of a volume hologram, and hence the single laser readout beam reconstructed all the images, of which the first four from the lower half plane are shown in Fig. 2c. Figure 2d presents the simultaneous reconstruction from the volume hologram (PR crystal) with the HeNe and Ar lasers.

References

(1) L. Pugliese and G. M. Morris, Opt. Lett. 15, 338 (1990).

(2) H. Kogelnick, Bell Sys. Tech. Jour. 48, 2909 (1969).

(3) J. P. Allebach and D. W. Sweeny, Proc. SPIE 884, 2 (1988).

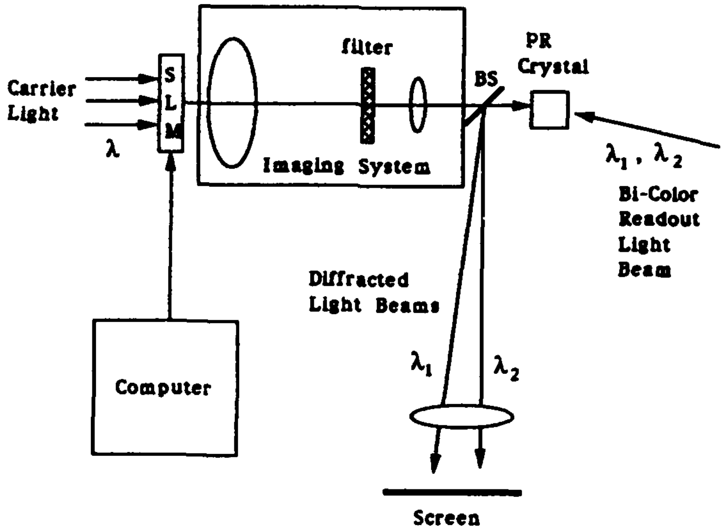


Fig. 1: The experimental setup for computer generation and recording of volume holograms.

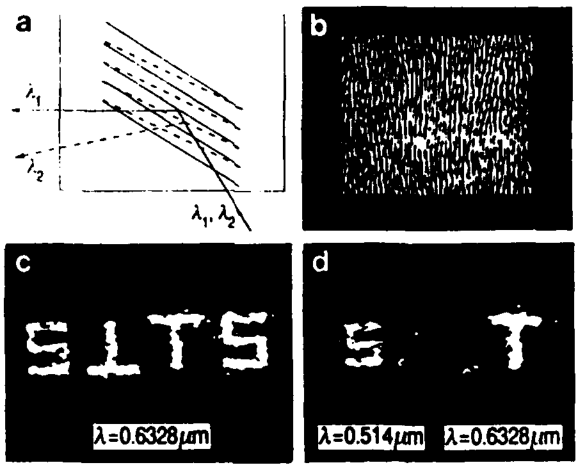


Fig. 2: (a) A detailed sketch demonstrating the reconstruction and its angular dependence upon the tilted wavelength-multiplexed volume holograms, (b) the binary computer-generated hologram, (c) the direct reconstruction from the SLM with an HeNe laser, and (d) the reconstructions from the volume holographic medium with the HeNe and Ar lasers simultaneously. The right image (T) is red, while the left image (S) is green.

Parallel Systems: 1

OWB 10:30am-12:10pm
Mesquite Room

Bernard H. Soffer, *Presider*
Hughes Research Laboratory

"Smart" Optical Interconnects for High Speed Photonic Computing

Peter S. Guilfoyle, Frederick F. Zeise and John M. Hessenbruch
OptiComp Corporation
PO Box 10779
Zephyr Cove, NV 89448
702-588-4176

Introduction

High speed, low power, "smart" DANE pixels are being developed to implement 2-D arrays of N-bit Boolean multiplications by DeMorgan's operation on wide-word FAN-INs coupled with N^4 global free space optical interconnects. "DANE" is an acronym that refers to the GaAs "smart" pixel which [1] detects light, [2] amplifies the result, [3] negates the result (inversion) and [4] emits the resultant Boolean value through the output laser.

DANE: Detection, Amplification, Negation, Emission

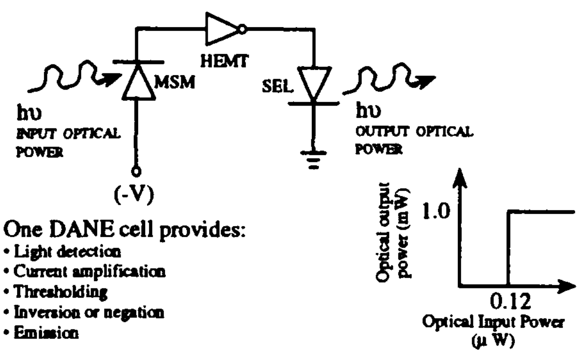


Figure 1: Simple example of a DANE cell

DANE switching technology is focused specifically towards the implementation of digital computing by augmenting the limitations of GaAs through the use of free space, "smart" optical interconnects. Compared to other smart cell approaches, the technological motivations are significantly different and unique because DANE device technology is architecturally driven. This high speed switch is readily implemented with conventional GaAs monolithic structures using processes compatible with E/D MESFET fabrication.

"Smart" Interconnect Algorithmic Efficiency

The ability to perform wide-word FAN-IN allows implementation of digital combinatorial algorithms which are currently not implemented with lower FAN-IN structures. Typically, the width of each gate has been limited by the FAN-IN. The use of photonic interconnect technology permits significantly more intensive computations at each gate delay level. Current algorithmic structures utilize extensive pipelining. Implementation of wide word FAN-IN allows a reduction in computational pipeline delays.

Figure 2 shows a simple example of a "smart" interconnect where the FAN-IN is N bits wide. When an arbitrary digital wide-word is represented in its complemented form, the detector acts as an OR gate, literally performing a Boolean summation on the wide-word. After electronic inversion and laser emission, the output light represents the N-bit AND product of the input bits. Consequently, a single gate delay with 64 bit AND gates is possible. These structures may be arbitrarily expanded to any digital function required such as wide-word addition, counting and floating point multiplication structures.

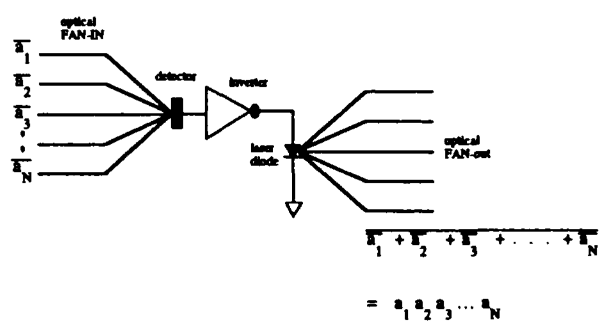


Figure 2: Simple example of "smart" interconnect which demonstrates DeMorgan's theorem on free space optical interconnects to achieve a wide input AND function

"Smart" *Global* (N^4) Optical Interconnects

Production of Minterms

The logical functions performed by the combination of DANE and "smart" interconnect are Shannon's minterms (functionals) at the photodetector array and the summation of minterms or complete instructions at the logical summation of the laser array. This is shown diagrammatically for 1D inputs and outputs in Figure 3.

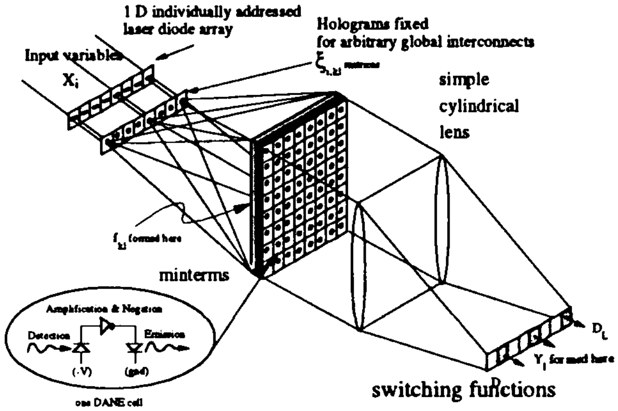


Figure 3: Global interconnects forming minterms, parallel output forming arbitrary switching functions

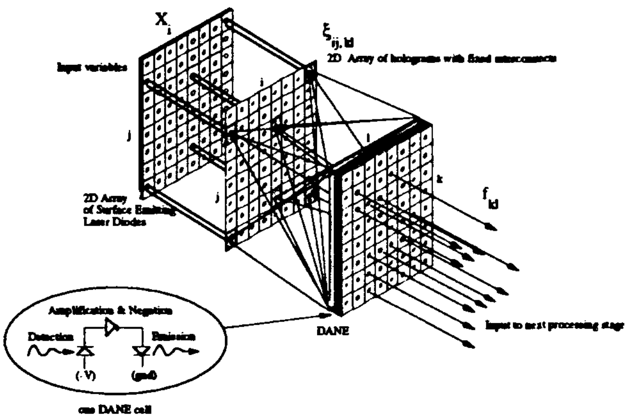


Figure 4: Single stage global free space "smart" interconnect module utilizing DANE optoelectronic computing devices forming arbitrary minterms.

Production of Arbitrary Switching Functions

The cylindrical lens shown in Figure 3, placed between the DANE and the output detector array, represents the most primitive (parallel) method of forming arbitrary switching functions. Summation of rows or columns of functionals at the DANE matrix output in the vertical (or horizontal) direction can be achieved with this simple cylindrical lens. The summed light is detected by the photodiode array producing Boolean summation by using a 0 - 1 threshold. The output Y_l of each photodiode thus represents a complete logical function or sum of products.

For parallel output, the logical function length is limited to **k** or l (depending on the lens orientation):

$$Y_l = \sum_{k=1}^{K}\left\{\prod_{i=1}^{N}\overline{x_i\xi_{i,kl}}\right\} \tag{1}$$

A spherical lens would provide summation of all selected minterms in both k and l dimensions, but only to one detector.

4-Dimensional Minterm Calculation

The optical scheme in Figure 3 can be generalized. As shown in Figure 4, it is possible to expand both the ξ input array to two dimensions and the linear holographic optical array to two dimensions. The holographic optical array is performing a 4 dimensional interconnect. The input is now written X_{ij}, the control mask is now $\xi_{ij,kl}$ and the output may be written $\mathbf{f}_{kl}$ where a two dimensional array of minterms is generated. Each minterm generated in the output array can consist of up to **i*j** Boolean variables. Equation 1 can thus be expanded as follows:

$$f_{kl} = \overline{\sum_{i=1}^{I}\sum_{j=1}^{J} x_{ij}\,\xi_{ij,kl}} = \prod_{i=1}^{I}\prod_{j=1}^{J}\overline{x_{ij}\,\xi_{ij,kl}} \tag{2}$$

4-Dimensional Switching Function Calculation

A second holographic optical interconnect element (HOIE), as shown in Figure 5 has the capability (subject to implementation constraints) of providing the summation over K * L minterms onto a detector plane. The summations for each individual detector are individually specified by the second HOIE. This true global interconnect is described by the equation below where KL and MN are the two dimensions of both DANE matrices respectively, and IJ is the number of independent input signals.

$$Y_{mn} = \sum_{k=1}^{K}\sum_{l=1}^{L}\left\{\prod_{i=1}^{I}\prod_{j=1}^{J}\overline{x_{ij}\,\xi_{ij,kl}}\right\}\eta_{kl,mn} \tag{3}$$

This architecture allows implementation of a larger FAN-IN which allows more than K or L minterms to be summed at each detector. For global output as shown in Figure 5, the GIBP is dependent only on the interconnects and the input bandwidth.

Consequently, for an I*J input array DANE matrix and K*L holographic matrix, up to N^4 interconnects may be accessed our global interconnect scheme. The theoretical GIBP of BN^4 represents the maximum possible complexity and the maximum number of realizable Boolean switching functions.

Applications

Free Space Optical Bus (3:8 MUX Decoder)

The "smart" optical interconnect technology implements a free space n bit decode using a single detector. Active signal routing can also be accomplished between these same 2^n destinations by using a bit serial protocol as in the

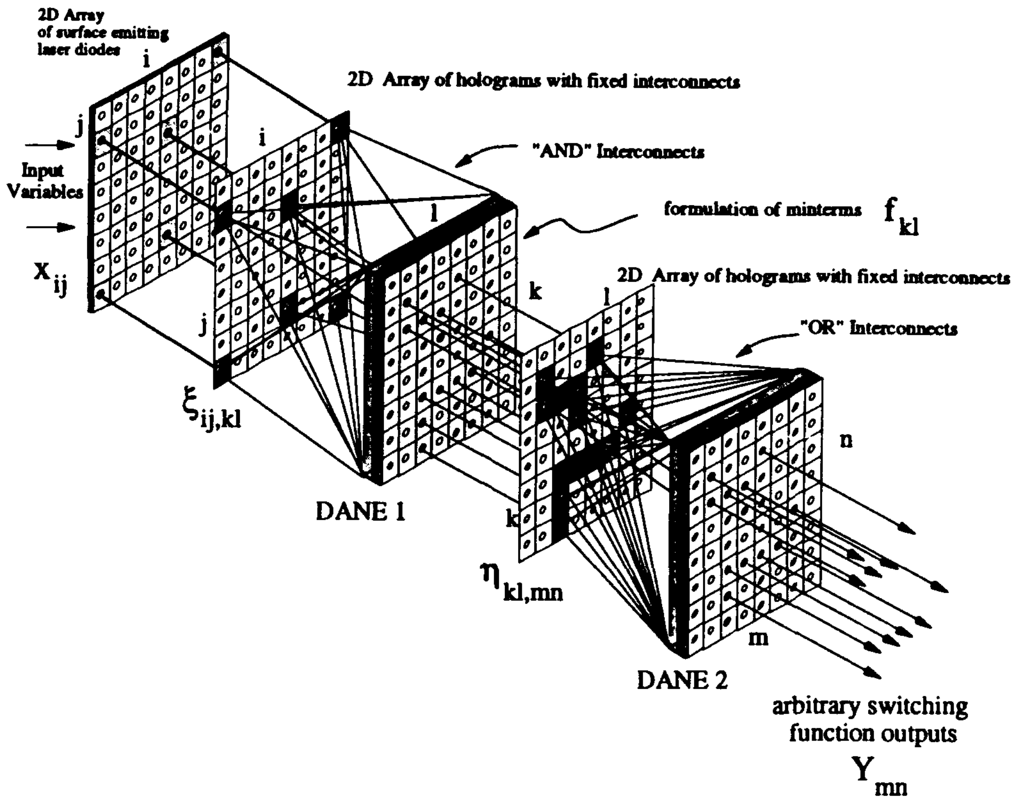


Figure 5: Two stages of the HOIE / DANE process allow for a complete set of arbitrary switching functions to be implemented through global free space smart optical interconnects

routed signal. The address need only be generated in dual rail form and utilize a control mask computer generated hologram (CGH) to implement the connection between the address bits and the specific address for each detector. The techniques which are used to implement the 3:8 decoder illustrate the principles behind the "smart" interconnect technology (Figure 6).

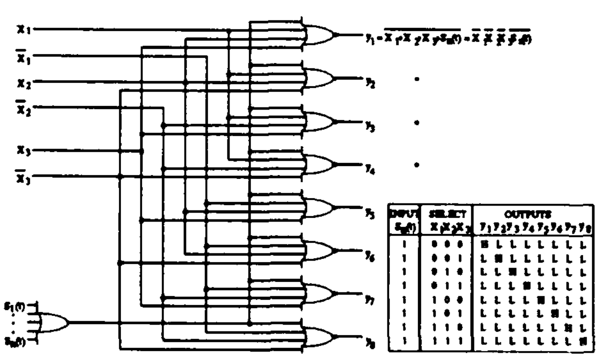


Figure 6: Digital logic representation of 3:8 MUX of a data router (simple BUS structure)

The digital circuit shown in Figure 6 has a 3 bit address field input represented in dual rail format, ie. x_1, x_2, and x_3, and their complements. The circuit, which has eight outputs: y_1 through y_8. is used to route a signal $s_n(t)$ to a given location. As the truth table in Figure 6 shows, the signal goes to one of eight locations depending on the address. This implementation utilizes the available signal sources with an efficiency of 50%.

MCM (Multichip Module) Data Routing

MCM interconnects are rapidly approaching their performance limits especially when future GaAs processor systems are considered. Furthermore, the number of I/O connections increase as the die size and MCM complexity become more apparent. Global optical interconnects using computer generated holograms (CGH) do not suffer the capacitive loading problems proportional to the speed and FAN-IN / FAN-OUT which limit electronic bus interconnects.

"Smart" interconnects may be applied to the MCM communication problem (Figure 7). "Simple"optical chip-to-chip level interconnects (point-to-point) have been suggested by several authors [1-4]. However, for clock distribution and point-to-point data transmission, these interconnects cannot perform processing functions. Using the global interconnect architecture, however, these

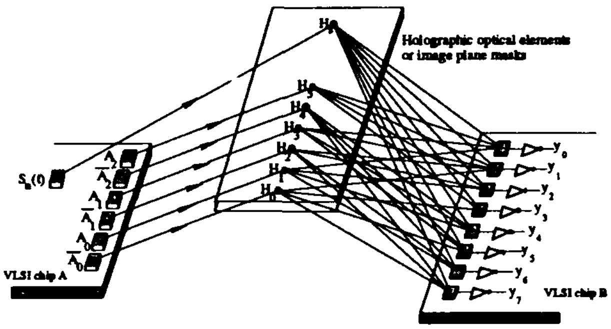


Figure 7: VLSI chip-to-chip level communications

"simple" interconnects can now be made "smart". For example, Figure 7 depicts the data routing of an arbitrary signal $s_n(t)$ from VLSI chip A to VLSI chip B by address application to the laser diode array on chip A. Previous schemes would require changing the masks for each fixed point-to-point interconnect. In this example, the mask never needs to be changed, and yet the data can still be routed through free space. Fourier transform reflective holograms can also be used to implement the mask and will support interconnects from and to the interior as well as the perimeter of the chip, thus easing the on chip data routing problem.

Algorithmic Efficiency and Projected Gate Interconnect Bandwidth Products (GIBP)

Table 1 demonstrates the interconnect efficiency for one clock wide-word addition. Efficiency is calculated by the number of interconnects used divided by the total number available for use. A "1" corresponds to a used interconnect; a "0" to a possible interconnect that the algorithm never uses. The gradual reduction of efficiency as a function of word width indicates that the global interconnect architecture has significant merit for wide-word processing [5].

Table 1: Total number of global free space interconnects for wideword widths

Bits	Total 1's	Total 1's and 0's	Efficiency
1	4	12	33.33
2	10	45	22.22%
3	20	112	17.86%
12	455	4225	10.77%
16	969	9537	10.16%
24	2925	30625	9.55%
32	6545	70785	9.25%
64	47905	545025	8.79%

The total gate interconnect in Table 1 is 47,905 for the 64-bit addition implementation. If a 5 nsec clock is used, the gate interconnect bandwidth product (GIBP) will be 9.581×10^{12}. The photodetector arrays allow an optical signal power per bit of 5000 photons, which corresponds to 1.2 fJ per gate. Theoretical signal power consumption at 10^{13} GIBP is only 100 mW. Total system power including DANE amplifier power when implemented should be no more than a factor of 10 above the total optical emission power. Typical silicon gates operate at 10 pJ which would correspond to a power consumption of 100 watts at a GIBP of 10^{13}.

Conclusions

The integration of GaAs DANE technology with global free space smart optical interconnects will:

•Provide "smart" pixel technology for arbitrary MCM interconnects.

•Increase the algorithmic functionality of GaAs ICs through free space optical logic.

•Augment GaAs IC interconnect limitations such as FAN-IN and FAN-OUT and delays.

•Decrease the power consumption of GaAs ICs (by up to two orders of magnitude).

Acknowledgements

OptiComp's research has been sponsored by the Strategic Defense Initiative Organization, through the Office of Naval Research.

References

[1] J. W. Goodman, F. Leonberger, H. Kung & R. Athale, "Optical Interconnects for VLSI Systems", Proc. IEEE Vol. 72, pp 850, 1984.

[2] J. W. Goodman editor, "International trends in Optics", 1991, Academic Press Inc., Boston.

[3] J. W. Goodman, "Optics as an Interconnect Technology" in H. H. Arsenault, T. Szoplik, B. Macukow, "Optical Processing and Computing", 1989 Academic Press Inc., Boston.

[4] M. Feldman, S. Esener, C. Guest & S. Lee, "Comparison Between Optical and Electrical Interconnects Based on Speed and Power Considerations", Applied Optics Vol. 27(9), May 1988.

[5] P. S. Guilfoyle, V. N. Morozov, "Potential Digital Optical Computer (DOC) III Architectures: the Next Generation", Proceedings of the SPIE, Vol. 1563-35, July, 1991.

Data/Knowledge Base Processing on an ODVM platform

Richard V. Stone and John Hessenbruch
OptiComp Corporation
P.O. Box 10779
Zephyr Cove, NV 89448
702-588-4176

Introduction

Through the utilization of an optical digital vector matrix (ODVM) processor, a mid range general purpose/ special purpose processor was designed and fabricated utilizing 8,192 optical interconnects. Designed to be retrofitted for a number of processing applications, this prototype was specified to execute 10^{12} binary operations/second while demonstrating near KT limit performance (5000 photon level threshold @ 10^{-16} BER). Tailored for the textural pattern recognition algorithm (TPRA), the ODVM processor has demonstrated the ability to scan > 20,000 pages of text/second. This machine addresses the interfacing and throughput issues which are commonly ignored in high speed parallel ODVM processing. Previous ODVM processors have been unable to demonstrate the full performance characteristics due to the limitations of the IO electronic feed .

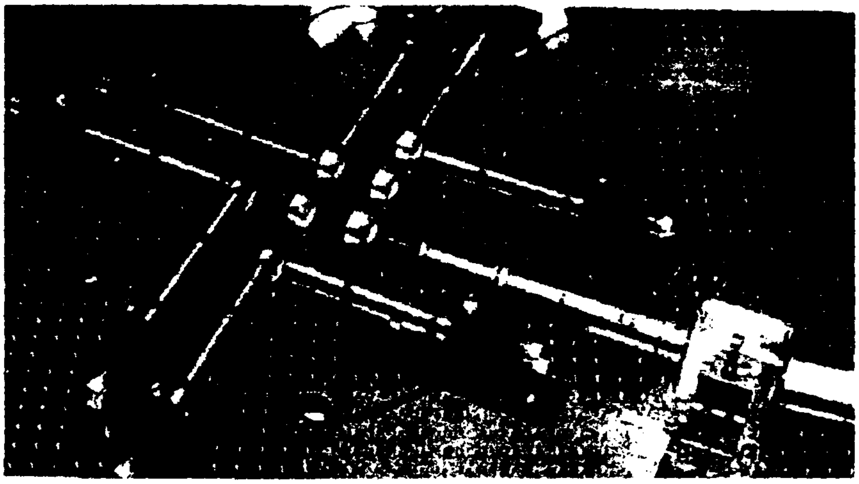


Figure 1: ODVM platform utilized for Data/knowledge base processing

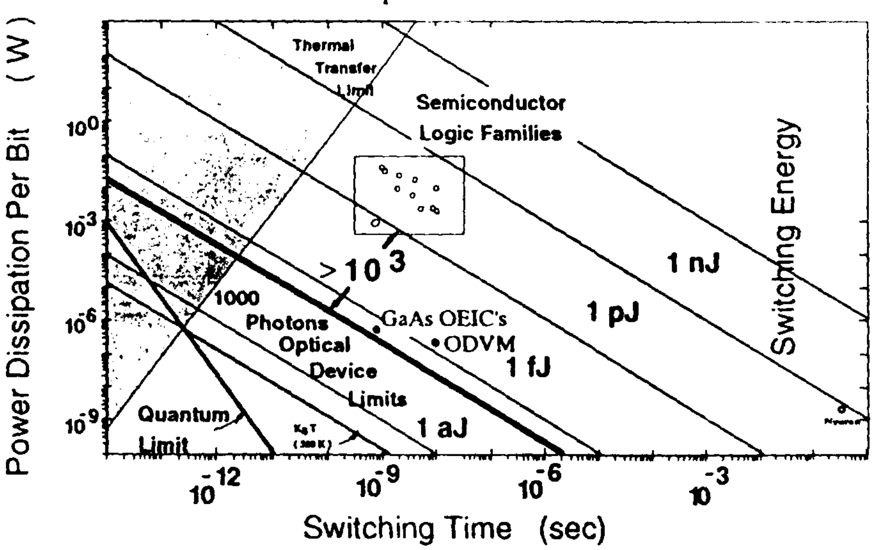


Figure 1b: Dissipative power vs. speed for digital logic families including ECL, optical digital vector matrix scheme and opto-electronic integrated circuit.

Optical Vector /Matrix configurations were developed in the 1960's and 1970's. The purpose of the development was to take advantage of two features which were analog based: high speed and parallel global algebraic operations [1]. The development of this trend still exists today without any significant breakthroughs in real world applications and implementations. However in the late 1980's it was realized that the vector matrix configurations could be readily adapted into the digital processing area [2,3]. ODVM processing offers the following advantages: 1) high-low threshold (Because low threshold can come significantly close to the KT limit lower system power consumption and increased gate densities are realized *see Figure 1b*). 2)1-N and N-1 interconnect configurations.

In order to achieve digital optical computing, three possible logic schemes are possible: residue logic, multilevel logic and binary logic [4]. It is our intention to demonstrate a binary based ODVM processor which is fully functional. By mapping Boolean logic primitives into multiple input planes and utilizing the inherent ability of the hardware to map the AND, OR, and invert functions, binary logic can be performed efficiently and can be mapped conveniently into a vector matrix form. In addition to the binary processor, the ODVM architecture is well suited to N bit equality detection for text searching of large data bases [5]. The original design of this platform takes an arbitrary vector 64 channels in length and convolves the vector with a 128 variable element matrix. The predominate issue of this ODVM platform becomes the I/O bottleneck concern. With the recent advancement in high speed digital electronics, the IO problem has been solved for this optical platform. With a sustainable clock rate of 100 MHz for data throughput, true parallel performance is demonstrated. This corresponds to a computational rate of $8x10^{11}$ binary operations/second. A demonstrated application for the ODVM processor is the wide word equality detection scheme.

N bit equality detection (NBED) text search

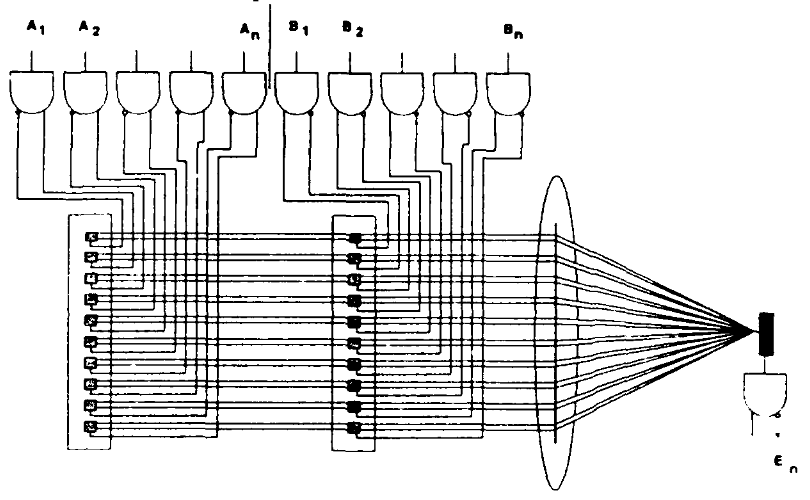
Most data knowledge base operations require searching a significant number of documents to retrieve and process only a few pages of data vectors. In this application, textural data is stored on optical disks, retrieved in parallel, and processed "on the fly" by the ODVM platform.
Full text search operations typical consist of the following operations as described in table 1:

TABLE 1. Full text search operations

1. Count the number of occurrences of word A in a document.
2. Search for word A in a document, paragraph or sentence.
3. Search for string A and/or string B in a document, paragraph or sentence
4. Search for words A and B with an arbitrary number of words in between.
5. Search for words A and B with exactly n words in between.
6. Search for the string X**Y (fixed length embedded "don t care").
7. Search for the patterns ?X or X? (variable length prefix or suffix "don t care")
8. Search for the patters X?Y (variable length embedded "don t care").
9. Count the number of sentences and/or paragraphs and/or words in a document

Following an approach taken by Guilfoyle [2], optical comparisons can be performed based on the exclusive-or (EX-OR) primitive using dual-rail logic. Two n-bit words A and B are equal if $A_i\overline{B_i} + \overline{A_i}B_i = 0$ for each pair of corresponding bits A_i and B_i. Because both the value of the bit and its complement are needed for the comparison, each n-bit word will be represented by 2n light

Figure 2: Optical comparison of two n- bit words based on the EX-OR primitive



beams (i.e. the 4-bit word 1011 will become 10-01-10-10). Using this method, a 00 combination corresponds to a "don't care" character, where as the 11 combination always produces a "not-equal" result. The coding scheme for the two logical values, 1 and 0, can be light and no light respectively.

As shown in Figure 2, the light beams are superimposed and focused on a single photodetector cell which performs the logical OR (or summation) of all the beams [3]. If no light is detected, the two words are equal, whereas any level of light intensity indicates that the two words differ in at least one bit. The output of the photodetector is electronic. Multiple word comparisons can be performed in parallel if a linear individually addressable laser array and a spatial light modulator are employed (Figure 3).

At any instance in time, the word at the i -th column of the IALDA is compared to the word at the i -th column of the SLM-2 array. The result, equal or not-equal, is recorded on the i -th cell of a row of photodetectors. This configuration allows for m comparisons of the n -bit words to occur simultaneously.

Figure 3: Vector matrix configuration for multiple word optical comparisons

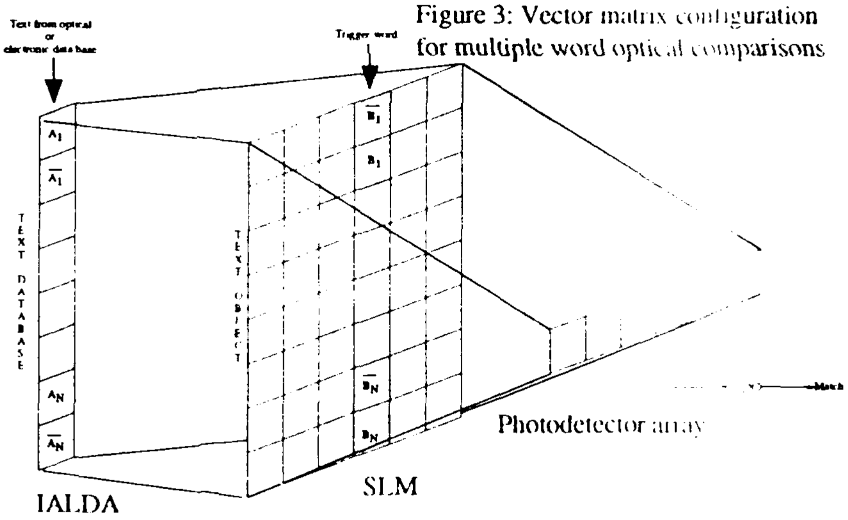


Figure 4 depicts the opto-mechanical layout that accepts electronic input into an individually addressable laser diode array bar. The output illumination from the IALDA is collimated through L1. A 10X image is formed at the output of L2. The first anamorphic relay (FAR,CL1,2,3) collimates and expands each of the 64 spots created at the output of L2 into 64 line images. The FAR provides uniform illumination across the 128 time bit aperture spatial light modulator. At the output of the SLM, is the second anamorphic relay which telecentrically images at 1.81X. Each of the 128 spatial time bits from a given input channel is imaged to one channel of a 128 channel avalanche photodiode array.

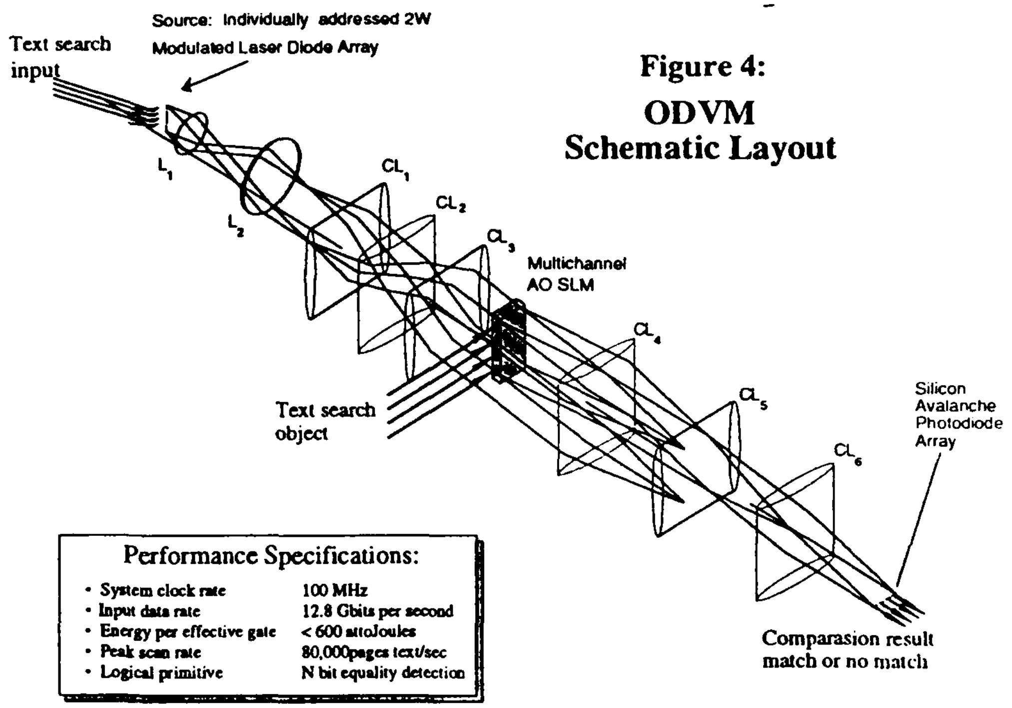


Figure 4:
ODVM
Schematic Layout

In the text search application, the search object which can be up to 128 bits long is input into the SLM. Data is then retrieved from either electronic or optical data bases and fed up to 32 lines into the individually modulated laser diode array bars. The lasers are stepped at 100 million characters per second. As data is fed into the lasers, it passes through the SLM. Thus 1.28x 10E10 16 bit character comparisons per second are demonstrated. The detector then records the position of the characters that match and record "a hit" and sends the corresponding "hit" signal to a query resolver.

Textural databases commonly contain documents of newspaper articles, case histories, technical articles etc. Current search approaches include full indexing of the data. However, as the database increases, the size of the index file becomes prohibitive. Using existing sophisticated computing equipment, the process is time consuming and expensive. This ODVM platform offers a solution to this problem because of the inherent bandwidth and speed offered by this system.

Acknowledgments

OptiComp's development of this platform has been supported by the Strategic Defense Initiative Organization and Rome Laboratory (USAF), through the Office of Naval Research.

References

[1] E.P. Mosca, RD. Griffin, F.P. Pursel, J.N. Lee, "Acoustooptical matrix-vector product processor: implementation issues", **Applied Optics**, Vol.28 No.18, September 15, 1989.

[2] P. S. Guilfoyle and W. Jackson Wiley, "Combinatorial Logic Based Optical Computing", Proceedings SPIE, Vol. 639-17, April 1986., **Applied Optics**, Vol. 27, May 1988, pp 1661-1673.

[3] P. S. Guilfoyle, "Programmable Optical Digital Computing", Proceedings of the 21st Annual Asilomar Conference on Signals, Systems and Computers, Nov. 3, 1987 and U.S. Patent 4,864,524 Sept. 5, 1989.

[4] K.S. Huang, BK Jenkins, A.A. Sawchuk, " Image algebra representation of parallel optical binary arithmetic", **Applied Optics**, Vol.28 No.6, March 15, 1989.

[5] P.A. Mitkas, P.G. Berra and P.S. Guilfoyle, "An Optical System for Full Text Search", Proceedings of 12th International Conf. on Research and Development in Information and Retrieval, Cambridge, MA, June 1989.

3-D Optical Memories

Sadik Esener
University of California-San Diego
Room R-007, 9500 Gilman Drive
La Jolla, CA 92093-0175

Summary not available at time of printing.

Comparison of Wavelength and Angle Multiplexed Holographic Memories

Geoffrey W. Burr, and Demetri Psaltis
Department of Electrical Engineering,
California Institute of Technology,
MS 116-81, Pasadena, Ca. 91125.
Ph (818) 356–4843

Kevin Curtis
Northrop B-2 Division, Pico Rivera, Ca. 90660, and
California Institute of Technology,
Department of Electrical Engineering,
MS 116-81, Pasadena, Ca. 91125.
Ph (818) 356–3888

Introduction

Multiple holograms can be recorded in a material by changing either the reference beam angle [1,2,3] (θ multiplexing) or the recording wavelength [4,5] (λ multiplexing). In the past, θ multiplexing has dominated primarily because convenient frequency-tunable light sources have not been available. With the advent of semiconductor and solid-state tunable lasers, there has been increased interest in λ multiplexing. Crosstalk SNR calculations [6,7] indicate that λ multiplexing appears to be advantageous since the SNR is independent of the total number of holograms for λ multiplexing, whereas it drops as one over the number of holograms for θ multiplexing. The purpose of this paper is to show that, in practice, the maximum number of holograms that can be stored with either method is not limited by crosstalk and therefore crosstalk should not be the sole, or even primary, basis on which the two methods are compared. We begin by showing that crosstalk in θ multiplexed storage can be suppressed by an order of magnitude by careful selection of the spacings of the reference beam angles. We then simulate the effect of noise in the setting of the reference beam angle on the crosstalk and show that unless this can be controlled accurately through feedback, this source of noise would be dominant. In the next step, we show that if L is the thickness and δ the resolution of the material, then there is an upper bound on the order L/δ from purely geometrical considerations for the number of holograms that can be superimposed for both λ and θ multiplexing . We show that when this upper bound is approached, λ and θ multiplexing yield comparable SNR. Finally, we require that the SNR for both schemes is the same and compare the number of holograms that can be stored in each case. This comparison leads us to the conclusion that the relative number of holograms that can be stored with a given SNR is determined by the tuning range of the laser and/or the spectral response of the material versus the numerical aperture of the optics used in the θ scanner and/or the readout optics.

Angular Multiplexing Strategies

Crosstalk between angularly multiplexed holograms arises from non–Bragg–matched readout of holograms stored at other reference angles by a given readout beam. This non–Bragg–matched readout has a sinc^2 dependence on the angular mismatch. We can eliminate crosstalk between a hologram and its neighbor by making the angular separation between two holograms equal to the distance to the first (or second) null of the sinc function. Unfortunately, the width of the sinc function varies as a function of reference beam angle. Therefore, if we select the same angular separation between all the holograms, crosstalk will result [6]. We will refer to this as the "uniform increment" strategy for angular multiplexing. We can improve the performance by making the separation between adjacent holograms variable and equal to the half-width of the sinc function for any two holograms. In Fig.1a, we plot the minimum SNR that results with the uniform increment as well as this "variable increment" strategy as a function of the hologram number. From Fig.1, we see that the variable increment strategy is always superior and for 2,000 holograms the improvement is approximately a factor of 10.

The SNR calculations in Fig.1a are calculated assuming the reference beam can be positioned with arbitrary accuracy at the designated angles. In practice, there will be uncertainty in the reference beam angle and these deviations can introduce additional crosstalk. Plotted in Fig.1b is the minimum SNR, averaged for 100 trials, of the reconstructed holograms obtained when a noisy beam deflector is used to set the angle of the reference. The simulations were done by adding Gaussian noise to the reference beam with a standard (angular) deviation equal to 3.3×10^{-6} radians, which is approximately consistent with the accuracy of our stepper motor when setting the angle of the reference beam using an optical system with a lens of focal length 300 mm. We observe two trends. The effect of positioning noise is most serious for relatively few holograms. This is reasonable since for a large number of holograms the deviations from the ideal zero-crosstalk positions become larger and the effect of the additional noise is less pronounced. The second observation is that the variable increment system still outperforms the uniform increment strategy even with positioning noise, particularly for a large number of holograms.

Comparison of Wavelength and Angular Multiplexing

For θ multiplexed Fourier-transform holograms in a 90° configuration with a uniform increment recording, the minimum SNR is given by [6]

$$SNR_{\theta} = \frac{2Lf}{\lambda dN}, \tag{1}$$

where L is the interaction length, f is focal length, λ is the wavelength, d is the full length of the output/image plane, and N is the number of holograms. For wavelength multiplexed

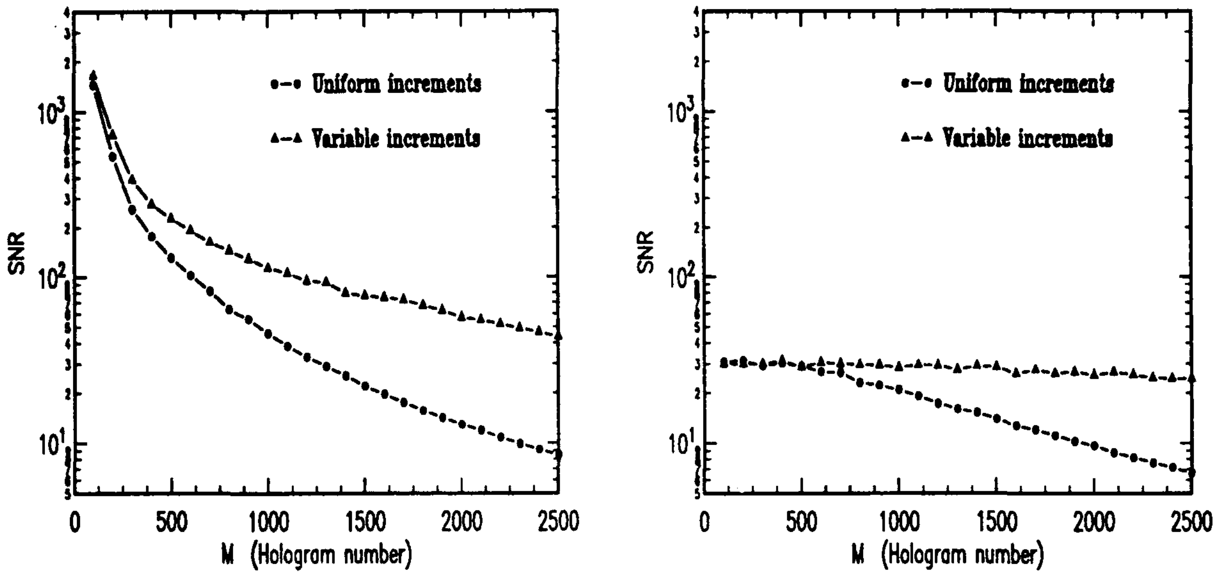


Figure 1: SNR for uniform and variable multiplexing schedule with (a) and without (b) random positioning errors

Fourier-transform holograms in a 180° configuration, the minimum SNR is given by

$$SNR_\lambda = \frac{f^2}{x_{max}^2 + y_{max}^2} \approx \frac{2f^2}{d^2}, \tag{2}$$

where f is focal length, x_{max} and y_{max} are half the length and width of the output/input plane, and d is the full length of the output/image plane. Therefore, for λ multiplexing, the SNR (for $N >$ a few holograms) is independent of the number of holograms.

We can divide Equation 2 by Equation 1 and get the ratio of the two SNR's. The number of holograms that can be stored is proportional to $\frac{L}{\delta}$ where δ is the resolution of the material. This fact immediately follows from a degrees of freedom argument and it is also relatively easy to prove for θ multiplexing. We will briefly show how it is also true for λ multiplexing. The largest optical frequency that we can use is $2c/\delta$ and the smallest is $2c/L$. The selectivity of the λ multiplexed holograms to a change in the optical frequency is $\delta\nu = c/L$. The number of holograms that can be λ multiplexed is therefore $c(\frac{1}{\delta} - \frac{1}{L})/(c/L) \approx \frac{L}{\delta}$. Using this upper bound for N, the ratio of SNRs becomes

$$\frac{SNR_\lambda}{SNR_\theta} = \frac{\lambda f N}{dL} \approx \frac{\lambda f}{d\delta}. \tag{3}$$

Typical parameters of $f/d = 10$ and $\lambda \approx \delta$ results in an SNR ratio ~ 10 when storing the maximum possible (geometrically limited) number of holograms. Notice that for thicker holograms, θ multiplexing becomes more favorable. Also, the SNR expression for θ multiplexing was for the uniform increment case. Switching to the variable increment method improves the SNR for θ multiplexing while there is no equivalent gain for λ multiplexing.

Another way to compare the two methods is by solving Equation 1 for N, and then using the SNR_λ from Equation 2 to obtain an expression for the number of holograms that can be θ multiplexed with the same SNR as that obtained by λ multiplexing:

$$N_\theta = \frac{Ld}{\lambda f}. \tag{4}$$

Using typical parameters ($L \approx 1$cm) results in $N_\theta \sim 10^3$. Also notice that increasing the thickness of the material increases the number of holograms that can be stored with the same SNR. However, in fairness, increasing L also increases the λ sensitivity which implies the number of holograms will also increase. Let $\Delta\nu$ be the spectral width of the light source and/or the spectral sensitivity of the material. Then the number of holograms that can be stored with λ multiplexing is

$$N_\lambda = \frac{\Delta\nu}{\delta\nu} = \frac{\Delta\nu L}{c} \tag{5}$$

We can now obtain an expression for the ratio of the number of holograms that can be stored with the same SNR for λ and θ multiplexing:

$$R = \frac{N_\theta}{N_\lambda} = \frac{\frac{d}{f}}{\frac{\Delta\nu}{\nu}} \tag{6}$$

Therefore, it is the fractional spectral bandwidth of the tunable laser versus the numerical aperture of the optical system that carries the stored images that determines which method can store more holograms in the same volume and with the same SNR.

References

1. D.L. Staebler, J.J. Amodei, and W. Phillips,"Multiple storage of thick phase holograms in $LiNbO_3$," VII Int. Quant. Elec. Conf., Montreal, May (1972).

2. F.H. Mok, M.C. Tackitt, and H.M. Stoll, "Storage of 500 high resolution holograms in a $LiNbO_3$ crystal," Opt. Lett. **16**, 605-607 (1991).

3. Demetri Psaltis,"Parallel optical memories," Byte, **17**, no. 9, 179 (1992).

4. H. Kogelnik,"Coupled wave theory for thick hologram gratings," Bell Syst. Tech. J. 48 (1969) 2909.

5. G.A. Rakujic, V. Leyva, and A. Yariv. "Optical data storage by using orthogonal wavelength-multiplexed volume holograms," Opt. Lett. **17**, 1471-73(1992).

6. C. Gu, J. Hong, I. McMichael, R. Saxena, and F.H. Mok, "Cross-Talk-Limited Storage Capacity of Volume Holographic Memory," J. OSA. A **9**, 1-6 (1992).

7. C. Gu, and J. Hong, " Capacity of wavelength multiplexed Holographic Memory,"OSA Annual Meeting, Albuqueque, Sept. (1992).

Optical Neural Networks

OWC 2:00pm-3:30pm
Mesquite Room

Kelvin H. Wagner, *Presider*
University of Colorado

Holographic Neural Networks: A Systems Perspective

Yuri Owechko

310-317-5839
owechko@csfvax.dnet.hac.com

Hughes Research Laboratories
3011 Malibu Canyon Rd.
Malibu, CA 90265

Many important problems in pattern recognition have no known algorithmic solution. These are problems at which biological organisms excel compared to machine-based approaches in which the rules are assumed *a priori*. Taking inspiration from living organisms, the neural network computation paradigm has been used by many workers to solve problems in which the underlying algorithm is unknown and the required transformations must be learned from examples. Neural networks have been shown to be universal approximators in that any function can be approximated to any desired degree by a sufficiently large neural network.[1] Optics, with its high connectivity, parallelism, and large storage capacity, has much to offer in terms of future implementations of large neural networks. In this paper I discuss some recent progress in holographic neural networks.

Neural network models of computation consist of many simple processing nodes or "neurons" which communicate with each other via weighted interconnections. Associated with each neuron is an activation level which is calculated from a weighted sum of the activity levels of other neurons.

$$y_i^{(n)} = f\left(x_i^{(n)}\right)$$
$$x_i^{(n)} = \sum_j w_{ij}^{(n)} y_j^{(n-1)}$$

Patterns which are input to the network through the bottom layer are transformed into patterns which represent the answer to the problem the network is trained to solve. Many neural net architectures have been designed and demonstrated for various computing tasks. Techniques have been developed to "train" or adjust the interconnection weights to solve problems in pattern recognition, vision, and robotic

[1] K. Hornik, M. Stinchcombe, and H. White,"Multi-Layer Feedforward Networks are Universal Approximators," Neural Networks 2, pp. 359-366 (1989).

control. Although the general problem of learning an arbitrary transformation for a completely random problem has been shown by Judd[2] to be NP-complete (therefore probably requiring exponential increases in learning time as the problem size increases), in practice specific problems in pattern recognition involve at least partially structured data for which the learning time can often be reduced.

Two important parameters which characterize a neural network are the number of neurons, N, and their connectivity, K, where K is the number of synapses or weights connected to a neuron. Both N and K should be as large as possible in general purpose neurocomputers. Pattern recognition problems involving natural data, especially vision, require large N to handle the large raw input data rate. Abu-Mostafa[3] has shown that K must be large for at least two reasons. First, since the neurons essentially implement K-input threshold functions, one K-input neuron with K associated weights is equivalent to order K^2 two-input neurons with $2K^2$ associated weights. Thus far fewer weights are required if the neurons have high fan-in and fan-out. Second, in order for a neural network to learn, the connectivity K must exceed the entropy H of the environment where H is the $\log_2$ of the number of input patterns typically generated by the environment. H increases as the randomness of the problem increases. Increasing N and K while maintaining computational parallelism is a daunting task for electronic architectures due to the 2-D nature of electronic interconnects. Most of the area on analog or digital electronic neuro-chips is taken up by the interconnects while implementing only moderate numbers of neurons.

Optical implementations of neural networks are attractive because of the large storage capacity and, most importantly, the parallel access and processing capabilities of optics. Optical architectures can exploit 3-D free space interconnects, allowing the input and output planes to be fully populated with highly interconnected neurons ($N=O(10^5)$). Moreover, an entire weight layer can be updated in one time step. The optical processors described in my talk use 3-D weight storage based on volume holography. The primary motivation for considering volume holograms as a storage medium for neural networks is the potential for extremely high storage capacity combined with fully parallel processing of the weights during both the learning and reading phases. Motivation for maximizing these parameters can be found in the potential application areas for neural networks.

Figure 1 is an adaption of a figure which originally appeared in the final report of the 1988 DARPA Neural Network Study. It shows the potential application areas mapped onto a 2-D space in which one axis is the storage required in terms of the number of weighted connections and the other axis is the processing rate required in connections per second. The corresponding estimated parameters of some biological organisms are included. A variety of electronic implementations, denoted by open

---

[2]S. Judd, "Learning in Networks is Hard," Proceedings of IEEE International Conference on Neural Networks, San Diego, 1987, p. II-685.

[3]Y. Abu-Mostafa, "Appendix D: Complexity in Neural Systems," in Analog VLSI and Neural Systems by C. Mead, Addison-Wesley, 1989.

squares, have also been added to Fig. 1. (These performance points were taken from an article by Alspector.[4]) The electronic implementations are, apart from the Sun and Cray computers, custom analog and digital chips. Some implement learning but some do not. For non-learning neural network chips, weight values must be learned off-line and subsequently loaded into the chip. I have added diagonal lines of constant network update time to the figure. The network update time is the time required for information to pass from the input of the network to the output via the interconnection weights. For a trained network, it represents the time needed to recognize and classify an input pattern. Such a plot is necessarily a highly folded and simplified projection of a high dimensional reality onto a low dimensional representation. For example, the degree of local vs global connectivity is ignored as well as the complexity of the neural network. Nevertheless, if we are careful certain trends can be deduced.

It is interesting to note that potential application areas in robotics, speech, and vision cluster around the 10 msec network update time line. These applications require large numbers of weights in order to achieve the complexity required to solve the problem, but modest solution times of a few milliseconds are satisfactory. The large number of weights, however, requires very high processing rates to achieve this update time. Significantly, biological organisms also cluster around the 10 msec update line, perhaps because they must also solve problems on the same time scale in speech and vision. With the exception of the general purpose computers, the electronic implementations have relatively modest storage capacity, although their processing rates are high. (Ignoring the fact that many of the chips do not have on-chip learning.) This is due to the 2-D nature of VLSI which limits the number of connections. They therefore appear most suited to the signal processing applications in which fast update times and modest storage are required.

In my view optical neurocomputers are complementary to specialized electronics in that the 3-D connectivity and parallelism of optics permit the implementation of very large networks with high processing rates and relatively modest network update times suitable for such applications as vision. In order to fulfill this role, optical neurocomputers should include, first, large numbers of neurons and weights; second, distortionless programmable mapping of a variety of neural network algorithms with no hardware reconfiguration; third, co-processor-type interfacing to a host computer; and fourth, hardware simplicity for low cost and compact packaging. A number of academic and industrial groups are developing holographic neural network systems and components. Some of the recent work will be reviewed in my oral presentation.

In 1992 a prototype laboratory version of a programmable holographic neural network computer was demonstrated at Hughes Research Laboratories. The Hughes Simulated Photorefractive Optical Neural Network (SPONN) is based on cascaded grating holography, a real-time holographic recording technique which greatly reduces several sources of distortion in holographic neural networks. In SPONN the weights are

---

[4] J. Alspector, "Parallel Implementations of Neural Networks: Electronics, Optics, Biology," Paper WB1, OSA Topical Meeting on Optical Computing, Salt Lake City, 1991.

distributed among many angularly and spatially multiplexed gratings. By utilizing a large fraction of possible grating k-vector values, SPONN reduces crosstalk, improves holographic image quality, and improves the utilization of the optical input device.

SPONN was built from available off-the-shelf components. Its present and projected future performances are indicated in Fig. 1. Up to 10^4 neurons, $2x10^7$ weights, and learning rates of $2x10^7$ connection updates per second have been achieved as of December, 1992. The potential exists for improvement factors of 10^3 in the number of weights and 10^6 in the processing rate. Perceptron, Bidirectional Associative Memory, and backpropagation networks with a single hidden layer have been successfully implemented by SPONN. The number of neurons, number of layers, and the neuron activation function can all be programmed without hardware adjustments. In addition, the system design requires only a single crystal, input spatial light modulator, and output detector regardless of the network configuration.

Cascaded grating holography has been demonstrated in photorefractive $BaTiO_3$ crystals using both visible light (514 nm) from an argon laser and infrared light (830 nm) from a laser diode. The infrared experiments are important because very compact systems can be built using laser diode light sources. Although the present laboratory system occupies a relatively large volume, packaging concepts based on laser diodes in which the entire system is contained within a volume smaller than a shoe box are under study.

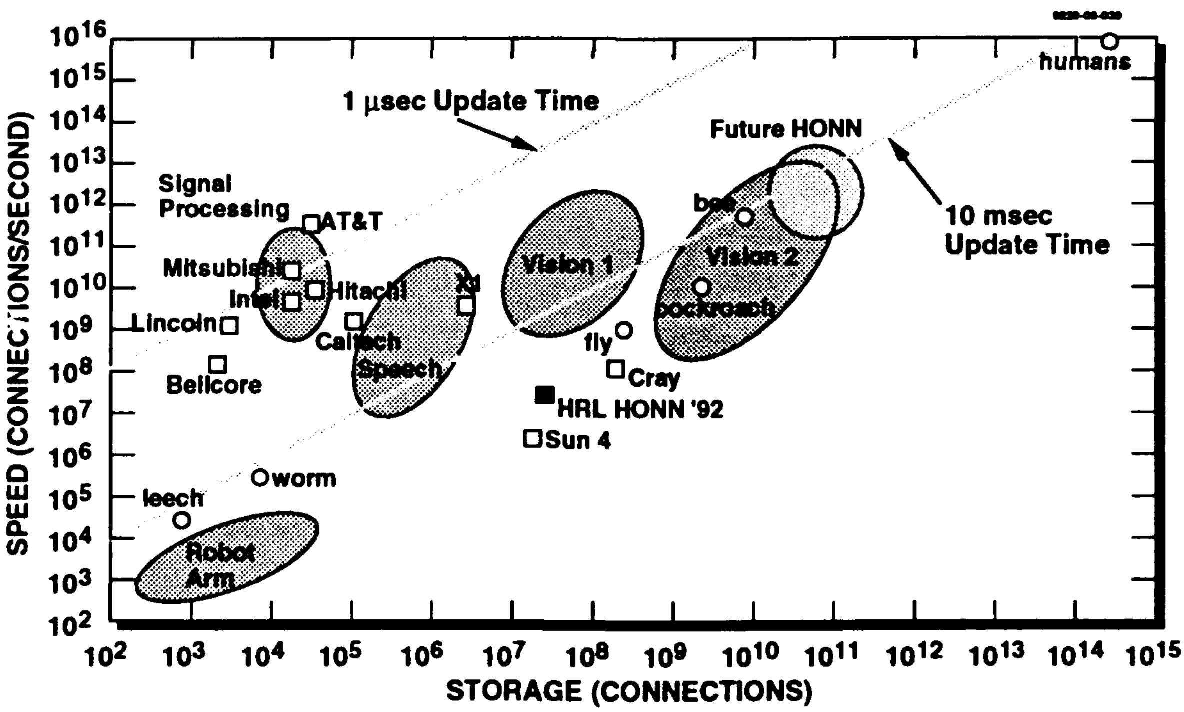


Fig. 1. Performance parameters of neural network implementations (electronic--open squares, Hughes optical--filled square, biological--open circles) compared to application requirements.

Second Order Neural Network Implementation Using Asymmetric Fabry-Perot Modulators.

Andrew Jennings, Brian Kelly, and John Hegarty,
Optronics Ireland Research Centre, Trinity College,
Dublin 2, Ireland.
Paul Horan,
Hitachi Dublin Laboratory, O'Reilly Institute,Trinity College,
Dublin 2, Ireland.

Introduction

Many optical neural network implementations using vector/matrix multipliers are based on first order networks in which each input neuron is connected to an output via a weighted interconnection. However, many interesting problems are of higher order, for instance translation invariance which is a second order problem [1]. One approach to solving such problems is to employ second order neural networks, with an advantage that the required training algorithms are much simplified compared with cascaded first order networks. This is done at the expense of increased complexity of interconnection, making the second order algorithm particularly well suited to optical implementation. Translation invariance has been studied by Giles and Maxwell [2] and has been implemented optically using Liquid Crystal Display (LCD) technology [3]. We now report on the implementation of a second order classifier network incorporating translation invariance, in which asymmetric Fabry-Perot modulator (AFPM) arrays based on semiconductor quantum wells are being used as optical input devices, to calculate an auto-correlation matrix and to implement weighted interconnects. The use of AFPM devices in the network allows for higher speed of operation and marks the first application of AFPM's.

Second Order Neural Networks.

A discussion of second order neural networks is given in reference 4, with a detailed discussion of the particular algorithm implemented in this work given in reference 3. A summary is given below. In a network consisting only of second-order interconnections, the output y_i is given by

$$y_i = f_{Th}\left(\sum_{j=1}^{N} \sum_{k=j+1}^{N} w_{ijk}\, x_j\, x_k \right) \qquad \text{Equ. 1}$$

where f_{Th} is a thresholding function, w_{ijk} are weighted interconnects and x_j and x_k are the terms in a one dimensional input vector. The thresholding function is usually a simple step function, y=1 for x>0, y=0 for x<0. However, during the training phase of the network a region of y=0.5 for $-\delta < x < \delta$ is defined, giving a region of "don't know" to the thresholding as a first approximation to a sigmoid. The product $x_j x_k$ correlates all inputs, yielding the auto-correlation matrix which is symmetric about the main diagonal. In an optical implementation it is possible to calculate all the terms in parallel [5].

During training, the network begins with small random weights and the weights are updated according to a simple perceptron rule

$$\Delta w_{ijk} = (t_i - y_i)\, x_j\, x_k \qquad \text{Equ. 2}$$

where Δw_{ijk} is the change in weight w_{ijk}, t_i is the desired training output and y_i is the output given by Equ. 1. This procedure is iteratively applied until the required training outputs are produced, at which stage no further changes will occur. Translation invariance requires that the network be transparent to translated versions of the training pattern and is implemented if w_{ijk} is the same for all j and k such that k-j= constant [2] The invariance condition leads to all terms on any one diagonal of the auto-correlation matrix being multiplied by the same weighting, greatly simplifying the implementation of weights [3].

Optical Implementation

In implementing the network, AFPM arrays were used as both optical input devices and weighted interconnects. The devices, grown epitaxially on a GaAs substrate, consisted of an optical cavity formed by two mirrors of reflectivities 32% and 98%. The absorption of the intracavity medium, 70 GaAs/AlGaAs quantum wells, could be modulated through the quantum confined Stark effect, so modulating the reflectivity of the device. Contrast ratios of ~22:1 and insertion loss of ~1.8dB were demonstrated with good uniformity across arrays and also between arrays [6]. Linear arrays of 21 "finger" patterned devices, each device 80µm wide and 2.5mm long, were fabricated with the device pattern designed to suit the architecture being implemented.

The optical system for the complete network is shown in figure 1. The input device array (Array 1) was illuminated with an expanded and collimated beam from a Ti:Sapphire laser tuned to 857nm, the optimum wavelength of operation of the devices. Each device was electrically addressed using the digital output ports of a computer. The image reflected from the devices contained the binary vector being presented to the network, as shown in the example of figure 2(a). In the working system, the device array was mounted at 45° to facilitate calculation of the auto-correlation matrix $x_j x_k$. This calculation was achieved by performing a double pass on the input device, with an effective 90° rotation of the image between passes. The retro-reflecting prism and imaging lens 1 shown in figure 1 performed this function [5]. On the second pass, the device was illuminated with an image similar to that shown in figure 2(a), but with the incident image rotated. After the second pass the resulting image was therefore the pixeled image shown in figure 2(b), which is the auto-correlation matrix of the input vector.

The weighted interconnects w_{ijk} were implemented in this system using an array of AFPM devices in gray scale operation (Array 2 in figure 1). The weighting device array had to be electrically addressed using analogue outputs from two 8 bit, multiplexed and latching D/A converters. With the D/A signals controlled from a computer, the addressing scheme was practical and still allowed for 'on-line' learning in the network. The simplification of the weighting scheme, as a result of the translation invariance, allowed a device array identical to the input devices to be used for the weighted interconnects, since the same weighting factor was required along any diagonal of the auto-correlation matrix. The symmetry of the auto-correlation matrix also leads to a 2-fold degeneracy in the matrix, so that only one half of the matrix was used in the weighting scheme. To implement positive and negative weights, the weighting devices were time mutiplexed, displaying alternately the positive and negative weights. A simple lens system imaged one half of the auto-correlation matrix on to the weighting arrray with a magnification of √2. This array was oriented with the "finger" modulators running parallel to the main diagonal of the matrix, and so applying a constant weighting factor to terms along a given line. Figure 2(c) shows a photograph taken from

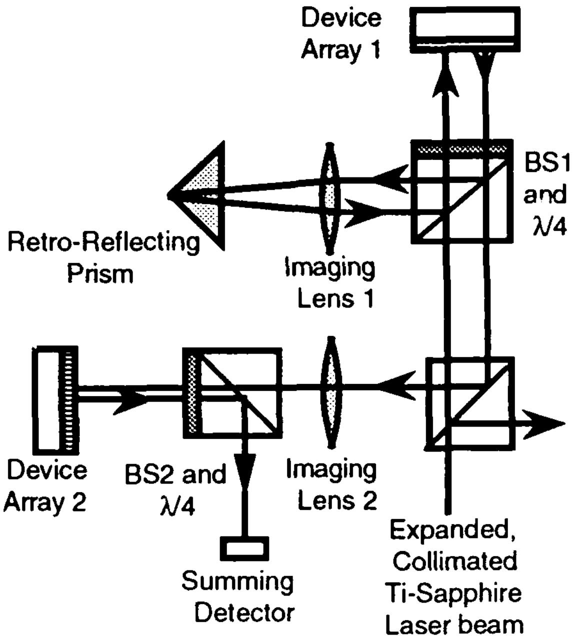


Figure 1. Schematic of optical network.

the weights devices when all the terms in the incident auto-correlation matrix are binary '1'. In this image, the weight strengths decrease from a maximum at the bottom most device to a minimum value at the 8^{th} device, indicating the gray scale signals being reflected off the weights. The top 4 devices in the image are at the maximum values.

The light from the weights was focused onto a single detector, summing the positive and negative weighted signals. The beam was chopped, allowing lock-in amplification to be used and the signal was read by a PC and thresholded.

Network Operation.

As a network with only one output neuron, this network could be trained to discriminate between two input vectors. The threshold function at the training phase of operation was optimised, so that a value of δ=0.02 volts was used with typical measured signals of 0.2 volts, and a weight capping value

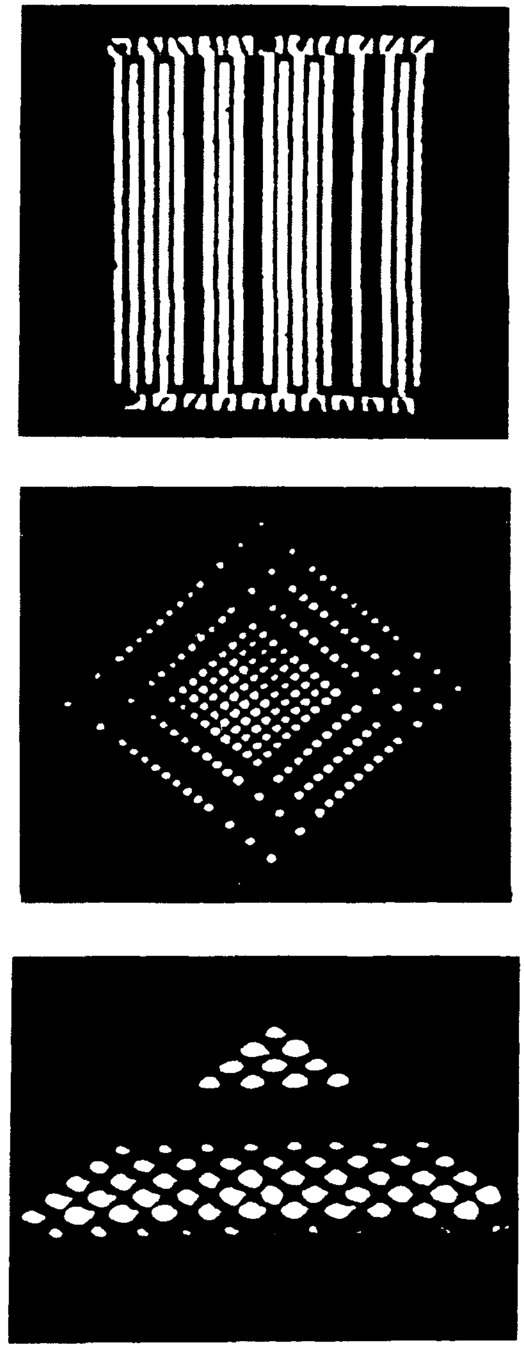


Figure 2: Photograph of

(a) image reflected off input devices.

(b) calculated auto-correlation matrix.

(c) image reflected off weighted interconnects demonstrating gray scale operation.

of 10 was used. The network was trained to discriminate between pairs of input vectors, typically converging to a set of weights after 6 iterations of the weights. The vectors were then subsquently presented to the network to monitor the recognition rate as the noise content of the input vector was increased. Figure 4 shows a typical response curve measured from the network as the noise content was increased to 50%, the recognition rate decreasing as the noise content increases. The recognition rate decreased from 100% recognition with 0% noise to approximately 40% recognition with 50% noise, demonstrating the 'on line' learning capabilities of the network and its noise tolerance in pattern recognition

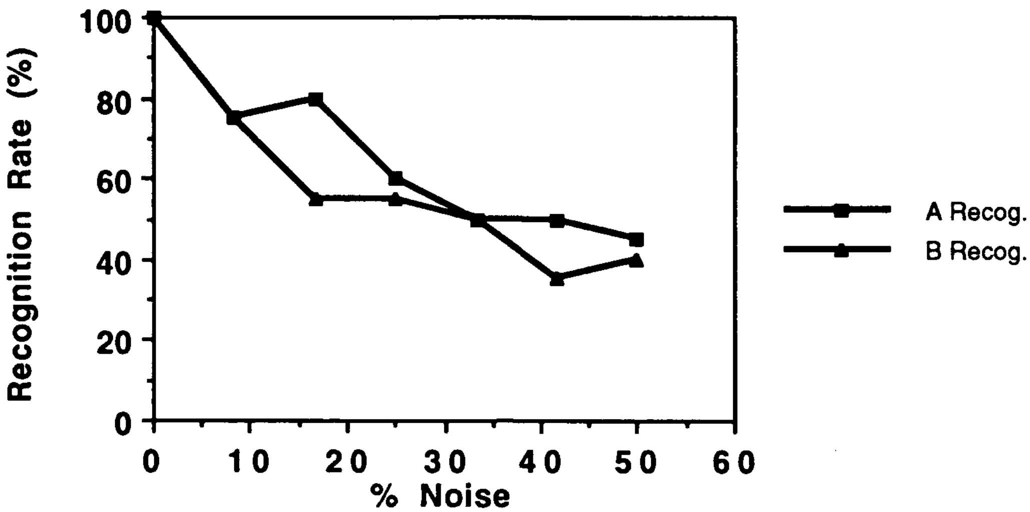


Figure 4 Recognition rate by network of trained pair with increasing noise at input.

Conclusion.

We have successfully implemented a second order neural network algorithm, incorporating translation invariance, in an optical system in which asymmetric Fabry-Perot modulator devices are employed as optical input devices and weighted interconnects. This marks the first application of AFPM devices in an optical system. The implementation allows for 'on-line' learning and shows good noise tolerance in recognition.

References.

1. M. A. Minsky and S. A. Papert, "*Perceptrons*" (MIT Press, Cambridge Mass., 1969).
2. C. L. Giles and T. Maxwell, "Learning, invariance, and generalization in high-order neural networks", App. Opt. **26**, 4972-4978 (1987).
3. P. Horan, A. Jennings, B. Kelly and J. Hegarty."Optical Implementation of a Second-Order Translation Invariant Network Algorithm", To be published in special issue on *Optical Neural Networks* App. Opt.
4. D. Psaltis, C. H. Park and J. Hong, "Higher order associative memories and their optical implementations" Neural Networks **1**,149-163 (1988).
5. A. Jennings, P. Horan and J. Hegarty. "Optical vector auto-correlation using asymmetric Fabry-Perot devices", Proc. of Optical Computing conference 1992, Minsk June 29th-July 3rd 1992.
6. A. Jennings, P. Horan, B. Kelly and J. Hegarty.
"Asymmetric Fabry-Perot Device Arrays with Low Insertion Loss and High Uniformity", Photonics Tech. Lett. **4**, (8), pp858-860, 1992

Content Addressable Network Implementations

Stephen A. Brodsky and Clark C. Guest
Department of Electrical and Computer Engineering
University of California, San Diego
La Jolla, CA 92093
(619) 534-6549

Content Addressable Networks are a family of neural networks designed for efficient implementation in optoelectronics and VLSI. Three CAN systems have been constructed for pattern classification, employing the supervised, self-organized, and tutored algorithm variations. The experimental systems use planes of parallel binary optical computations for both learning and recall. Experimental results are presented to demonstrate the fault tolerance of the supervised CAN algorithms. The supervised CAN network was able to learn around optical errors resulting from noise, stuck pixels, and failing pixels, even when these errors caused imperfect application of the learning algorithm.

One of the goals of Artificial Neural Network (ANN) implementations is the construction of highly efficient computational systems. Many ANN models include significant computational requirements[1] naturally expressed as highly parallel operations. However, the high precision computation and storage required by these models limits fully parallel implementations. Practical fully parallel ANN systems require both high speed and high density implementations.

The Content Addressable Network (CAN) family of ANNs overcome this limitation by reducing essential network computation and storage complexity, allowing construction of efficient systems employing fine grain parallelism. CAN networks incorporate a unique learning algorithm enabling development of networks based on binary arrays and planes of Boolean operations. This reduction allows construction of ANNs with fully parallel, high speed and high density components.

Three algorithms may be used to train CAN networks as pattern classifiers. The supervised algorithm trains a CAN network to learn the classification mapping between input and target pattern pairs. The self-organized model groups patterns within a user-defined level of similarity. Finally, the tutored algorithm allows a teacher to aid in defining pattern classes using 'yes' or 'no' responses.

The three CAN learning algorithms have been implemented on an experimental system. Two Liquid Crystal Televisions (LCTVs) were used as Spatial Light Modulators (SLMs) for optical representation of the network and data. The system performed optical computations using planes of binary operations, representing the CAN operations for recall and supervised learning. The supervised CAN algorithm was also tested against various optical errors as a demonstration of fault tolerance.

The supervised CAN learning algorithm[2] solves problems in a manner analogous to the popular Backpropagation[3] training algorithm. Both supervised CAN and Backpropagation may be used to adapt a multilayer feedforward network to perform a specified classification mapping, and both perform corrections to arrays of connection weights using an error reduction procedure. The key difference is that the CAN algorithm uses an intrinsically binary learning procedure to adjust arrays of binary connection weights. Supervised CAN is not a quantized version of Backpropagation, but a new method for learning on binary weights using the substantially reduced computation complexity of binary operations.

The following summarizes the supervised CAN network and learning algorithm. A more detailed explanation is available in Reference 2. An extension of the algorithm allowing learning for multi-valued analog problems is described in Reference 4. The implemented network consists of two feedforward mapping layers trained on binary pattern vector pairs. The binary weight layers connect the input vector to the hidden vector, and the hidden to the output. The hidden vector contains the network's internal representation, allowing a two-layer network to solve all mapping problems[5]. The output vector is compared to the target vector from the training set. Differences between the output and target are encoded in a binary error vector δ. This error vector is then used to make corrections to the layers' weights and thresholds. After multiple presentations of the training set, the network converges to the desired mapping.

Since the network and the training set contain only discrete components, the network obtains exact solutions. The CAN algorithm converges quickly, typically requiring far fewer presentations of the training set to reach zero error than an equivalently sized Backpropagation network would take to reach a learning plateau. Also, since the network only makes changes in response to error, overtraining is not possible.

Each CAN layer contains an array of binary connection weights, arranged into rows of neurons, or decision units. Each row unit has a discrete threshold, with a range equal to the number of weights in the row. When an input pattern is projected to the layer, the Hamming distance between the input pattern and each row of weights is computed using a summation of Exclusive-Or ($\oplus$) operations. The distance for each row unit is compared to an adaptive row threshold. If the summation equals or exceeds the threshold then the unit output is 1, otherwise it is 0.

Both the weights and the threshold are changed during learning. To briefly describe the learning procedure, imagine a single decision unit with a particular connection weight W and input I. The unit makes a decision based on whether the summation equals or exceeds the threshold. If the output O of the unit is in error, the balance between summation and threshold is incorrect. A correction is made by changing the relative size of one of the components to produce the correct balance. For a threshold, a correction means an increment or decrement; corrected weights are complemented. A weight is eligible for correction when the following conditions are true: The output of the unit is in error ($\delta = 1$), and the connection contribution ($W \oplus I$) is the same as the output O. This computation for whether a weight is eligible for correction may be done in parallel for all the weights in the whole layer, and is performed optically in these implementations.

The self-organized CAN network classifies binary patterns based on Hamming distance. The user controls the granularity of the classification by a tolerance parameter, analogous to the vigilance parameter of the ART 2 network[6]. A large tolerance, or low vigilance, allows wider variation between members of the same class. Correspondingly, low tolerance results in greater numbers of tighter classes. Each class also has an individual tolerance, allowing some classes to be more exclusive than others. For example, the tolerance could be based on the number of 1 pixels in a pattern, allowing classes with few 1 pixels to draw tighter distinctions.

The self-organized CAN consists of a single CAN layer. Class templates are stored in a non-distributed representation as rows in the layer. A query compares the input pattern with all class rows in parallel by computing the Hamming distance for each class, and the output represents the number of matching rows. When there are no matches, the thresholds are decremented toward the user defined tolerance parameter. When a single output match occurs, the query pattern is a member of the matched class. Otherwise, a new class is formed based on the query pattern. The thresholds are reset for the next query.

The tutored CAN network performs teacher directed self-organization. The network classifies binary patterns based on Hamming distance from stored class templates. This default classification may be modified by the 'yes' or 'no' responses from the teacher. 'Yes' responses allow the network to perform in a self-organizing manner, while 'no' responses direct the network to make successive attempts to associate the query pattern. Each 'no' response requires the network to provide its next best match until a correct ('yes') match occurs. The query pattern is then committed to the correct class.

The tutored CAN consists of two CAN layers, one for templates and one for template/class association. The first layer operates similarly to self-organized CAN, allowing a user-defined tolerance for the network and also for each class template. The first layer is augmented by an enable vector, which is used to suppress rejected matching templates. The second layer then directs a class template to its corresponding class member. This allows classes to have multiple templates

The three CAN algorithms have been implemented on common hardware. A single CAN layer was constructed from two LCTV SLMs between crossed polarizers, a detector array, and electronic control circuitry. The CAN layer was used consecutively as layers 1 and 2 for the supervised and tutored algorithms.

The optical system computed parallel arrays of the Boolean operations Exclusive-Or and AND. The computations were performed using polarization rotation. The Exclusive-Or was performed by rotating incoherent vertically polarized light 0 ° or 90° based on the SLM pixel values. AND was performed using 60° rotation. The data pixels were dual-rail encoded to compensate limitations of the LCTV internal control circuitry. All optical computations were fully parallel on arrays of pixel data.

The electronics consisted of an IBM AT with 2 image boards, a vidicon camera, and external video support circuitry. Electronics were used for detection, thresholding, summation, and sequencing. The speed of the network was limited by the PC bus's slow updates to the image boards. One board handled detector input, the other encoded both LCTV outputs. A previous implementation of the supervised CAN algorithm is reported in References 7 and 8.

Several two layer supervised CAN networks were trained using the experimental system. Two networks sizes were used, one with 25 decision units per layer, and one with 10 decision units per layer. Optical computations were performed for all the recall and learning operations for weights. (Because of limitations of the LCTV circuitry, the AND function was performed electronically for the 25 unit network, and optically for the 10 unit network.) Summation and threshold operations were performed electronically.

The 25 decision unit per layer network was trained to learn 2 pattern pairs in 22 presentations of the training set. The network is capable of learning up to 25 pattern pairs, but the speed of the PC portion of the network prevented longer training sessions. The system also learned the pattern set in the same number of presentations in the presence of a small amount of optical noise which produced approximately 1 error every 3 passes.

The 10 decision unit per layer network was also trained to learn 2 patterns. The network was trained with and without 3 blocked and 2 damaged pixels of 100 total. The network learned in 10 passes without obstruction, and in 28, 20 and 8 passes with the obstructions. The different number of passes is an indication of the inconsistency of the damaged pixels. In a different trial, when optical power was cut off temporarily during training, the system was able to recover and learn after optical power was restored.

The intentional errors introduced during optical computation affect both network recall and learning. Errors in computations for one layer are passed to the next during the recall and correction operations. The system is still tolerant under the supervised CAN training algorithm, even when optical faults cause imperfect network updates.

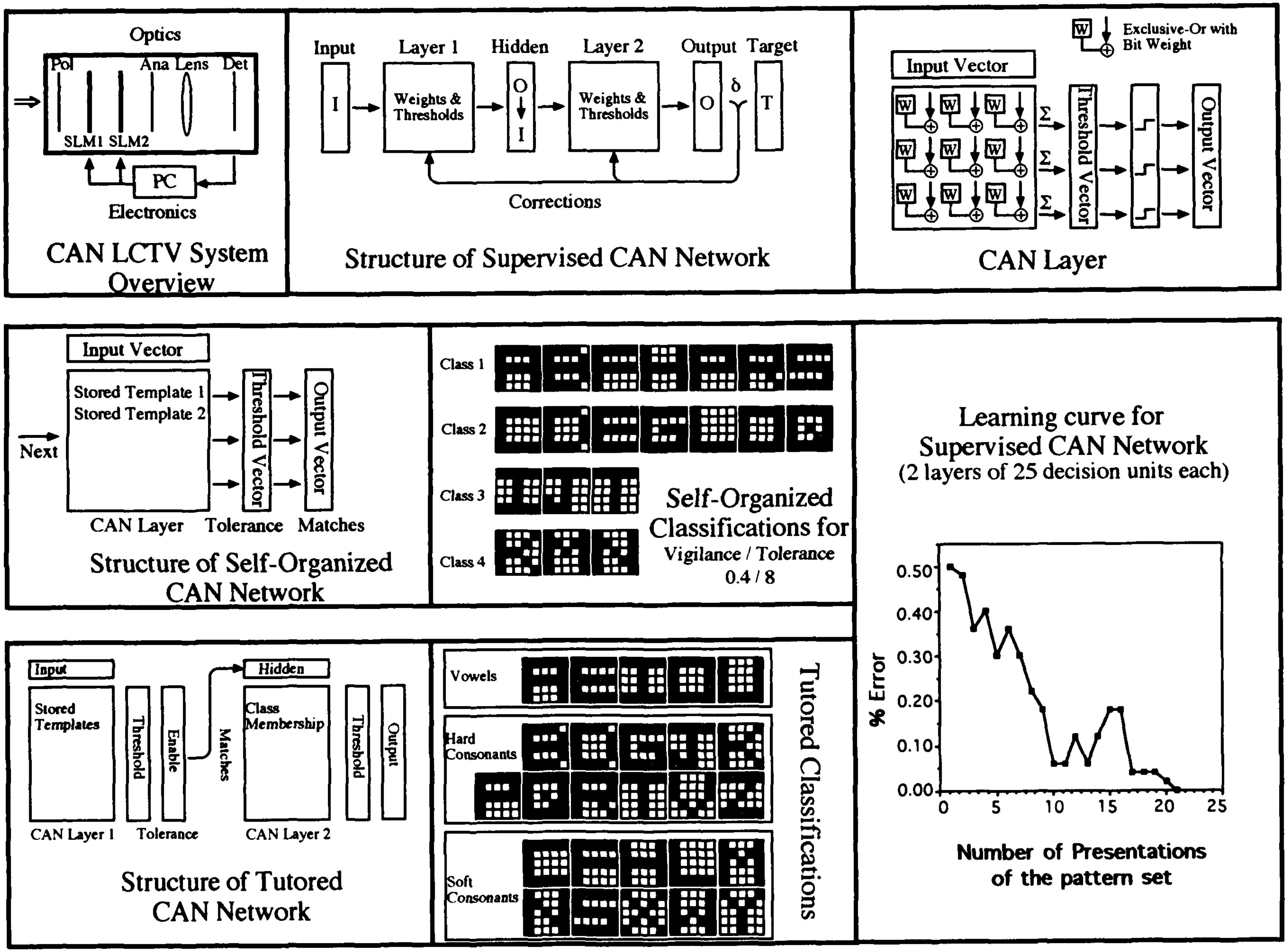
The experimental setup was re-used for the self-organized network. The system involved a single CAN layer with 25 decision units. The network was trained on 20 patterns, representing the first letters of the alphabet in the 5x5 font used for training the ART 2 network[6]. The network was trained at two vigilance/tolerance levels (0.4/8 and 0.6/5) producing class groupings of 4 and 10 classes, with average maximum Hamming distances per class of 7.0 and 1.8. For sake of comparison, the given ART 2 classifications (vigilance 0.5 and 0.8) produced groupings of 4 and 9 classes, with average maximum Hamming distances per class of 10.8 and 4.3. From the standpoint of maximum Hamming distance per class, the self-organizing CAN network produces more effective groupings.

The experimental system was also used for the tutored network. Two CAN layers with 25 decision units each were trained on the same font, this time including all 26 letters. When the system was allowed to classify along purely self-organizing principles (only 'yes' responses) the results were identical to the self-organized system. The system was also trained using both 'yes' and 'no' to divide the alphabet into 3 groups based on pronunciation groupings (vowel, hard or soft consonant), which are not related to the patterns' image.

The Content Addressable Network learning algorithms allow construction of classification networks by significant reduction of component complexity. Three members of the CAN algorithm family were implemented on common optoelectronic hardware, employing planes of parallel binary computations. The experimental supervised CAN algorithm used optical computations for both recall and learning, as well as demonstrated fault tolerance to optical errors from noise and damaged pixels. The self-organized CAN algorithm performed pattern classifications which compare favorably to the ART 2 network, and the tutored CAN system demonstrated optoelectronic operation of a new type of network. These implementations demonstrate the advantages of the highly parallel, binary based CAN algorithms in terms of reduced implementation complexity and intrinsic fault tolerance.

References

1. Hecht-Nielsen, R., Neurocomputing, Reading, Mass: Addison-Wesley, 1990.
2. Brodsky, S. A. and Guest, C. C., "Binary Backpropagation in Content Addressable Memory," International Joint Conference on Neural Networks (1990).
3. Rumelhart, D. E., McClelland, J. L., and the PDP Research Group, "Learning Internal Representations by Error Propagation," Parallel Distributed Processing, Explorations in the Microstructure of Cognition, Volume 1, Cambridge, Mass.: MIT Press (1986).
4. Brodsky, S. A. and Guest, C. C., "Content Addressable Networks for Initialization of Backpropagation," International Joint Conference on Neural Networks (1992).
5. White, H., "Connection non-parametric regression - multilayer feedforward networks can learn arbitrary mappings," Neural Networks, 3(5) 535-549 (1990).
6. Grossberg, S., "Nonlinear Neural Networks: Principles, Mechanisms, and Architectures," Neural Nets, 1(1) 17-61, (1988).
7. Brodsky, S. A. and Guest, C. C., "Optical Matrix-Vector Implementation of Binary Valued Backpropagation," Optical Society of America Conference on Optical Computing (1991).
8. Brodsky, S. A., Marsden G. C., and Guest, C. C., "Optical Matrix-Vector Implementation of the Content Addressable Network," Applied Optics (under review).


Optics
Pol
Ana Lens
Det
SLM1 SLM2
PC
Electronics
CAN LCTV System Overview
Input
Layer 1
Hidden
Layer 2
Output
Target
Weights & Thresholds
Weights & Thresholds
I
O
I
O
δ
T
Corrections
Structure of Supervised CAN Network
Exclusive-Or with Bit Weight
Input Vector
Threshold Vector
Output Vector
CAN Layer
Input Vector
Stored Template 1
Stored Template 2
Next
Threshold Vector
Output Vector
CAN Layer
Tolerance
Matches
Structure of Self-Organized CAN Network
Class 1
Class 2
Class 3
Class 4
Self-Organized Classifications for
Vigilance / Tolerance
0.4 / 8
Learning curve for
Supervised CAN Network
(2 layers of 25 decision units each)
% Error
0.50
0.40
0.30
0.20
0.10
0.00
0
5
10
15
20
25
Number of Presentations of the pattern set
Input
Stored Templates
Threshold
Enable
Matches
Hidden
Class Membership
Threshold
Output
CAN Layer 1
Tolerance
CAN Layer 2
Structure of Tutored CAN Network
Vowels
Hard Consonants
Soft Consonants
Tutored Classifications

Random Interconnections with Ground Glass for Optical TAG Neural Networks

Hyuek-Jae Lee, Soo-Young Lee, and Sang-Yung Shin
Department of Electrical Engineering
Korea Advanced Institute of Science and Technology
373-1 Kusong-dong, Yusong-gu
Taejon 305-701
Republic of Korea
Tel.: +82-42-869-3431, Fax: +82-42-869-3410, E-mail: sylee@eekaist.kaist.ac.kr

1. Introduction

After the first optical implementation of the 1-dimensional Hopfield model had been reported [1], extensive researches have been conducted for 2-dimensional input/out patterns.[2,3] Page-oriented holograms [4] may achieve fairly large fixed interconnections, while volume holograms [5] or lenslet arrays with spatial light modulators (SLMs) [6] achieve adaptive interconnections. However the volume hologram still requires further researches, especially on fixing and copying, and SLM resolution is major limiting factor for large scale implementation for the latter. Recently we had developed an adaptive neural network architecture, TAG (Training by Adaptive Gain), which utilizes fixed global interconnections and adaptive local gains.[7,8] Performance with both random and pre-trained interconnections were investigated. In the previous papers we had used page-oriented holograms for the fixed interconnections. In this paper we show possibility of using ground glass for random interconnections with much higher diffraction efficiency and interconnection density, and report small-scale optical implementation of a classifying neural network.

2. TAG Neural Network Architecture

Let's consider a single-layer neural network with $N \times N$ input neurons and $M \times M$ output neurons. For fully-connected adaptive neural networks one has M^2N^2 adaptive elements. In our original TAG model the interconnections are composed of N^2M^2 global fixed interconnections and N^2+M^2 local adaptive gain-controls. [7] For the modified TAG model the local adaptive interconnections were increased to $\alpha\,(N^2+M^2)$ for much better performance, where α denotes local connectivity. Fig.1 shows this architecture in a simple form. In mathematical notations output y_{ij} is represented as

$$y_{ij} = S(\hat{y}_{ij}) \ , \quad \hat{y}_{ij} = \sum_{k,l} v_{ij} T_{ijkl} w_{kl} x_{kl} \ , \tag{1}$$

where x_{kl} and y_{ij} denote activations of klth input neuron and ijth output neuron, respectively. The $v_{ij} T_{ijkl} w_{kl}$ denotes interconnections between x_{kl} and y_{ij}, which consists of fixed global interconnection (T_{ijkl}) and adaptive local gain-control (w_{kl} and v_{ij}). Here double indices are used for neurons to clarify 2-dimensional nature of the input/output patterns. $S(.)$ is a Sigmoid function.

We have adopted gradient-descent least-square-error minimization algorithm for the adaptive learning. The total error E is defined as

$$E = \frac{1}{2} \sum_{s} \sum_{i,j} (y_{ij}^s - t_{ij}^s)^2 \ , \tag{2}$$

where s is an index over classes (input-output pairs), y is the actual state of an output neuron, and t is its desired state. The partial derivatives may be obtained by chain rule as

$$\frac{\partial E}{\partial v_{ij}} = \sum_{s} \delta_{ij}^s \hat{y}_{ij}^s \,/\, v_{ij} \ , \tag{3}$$

and

$$\frac{\partial E}{\partial w_{kl}} = \sum_{s} \gamma_{kl}^s x_{kl}^s \,/\, w_{kl} \ , \tag{4}$$

where δ_{ij}^s and γ_{kl}^s are output and input errors, respectively, and defined as

$$\delta_{ij}^{s} = (y_{ij}^{s} - t_{ij}^{s}) S'(\hat{y}_{ij}^{s}), \quad (5)$$

$$\gamma_{kl}^{s} = \sum_{i} \delta_{ij}^{s} v_{ij} T_{ijj} w_{kl} . \quad (6)$$

It is worth noting that the input error γ_{kl}^{s} and may be calculated by back-propagation of output error, δ_{ij}^{s}. This error back-propagation allows us to extend this model to general multi-layer architectures.

The fixed interconnections may be randomly generated, or obtained from any learning algorithm for standard input/output patterns. For the former only one random interconnection hardware may be used for a wide variety of applications, which paves a way to practical implementations. For the latter the local gains may be trained to adapt to a specific user. Performance per adaptive elements of this TAG model is regarded as better than that of perceptron.[7]

3. Ground Glass for Random Interconnections

In globally connected single-layer neural networks such as perceptron and Hopfield model, SLM resolution directly limits achievable input neuron numbers multiplied by output neuron numbers. In our TAG model only sum of input neuron and output neuron numbers are limited by SLM resolution. However there still exists a problem for large-scale implementation of fixed global interconnections. The page-oriented holograms used in Ref. [8] show limited diffraction efficiency and require high-precision equipments for very high-density interconnections.

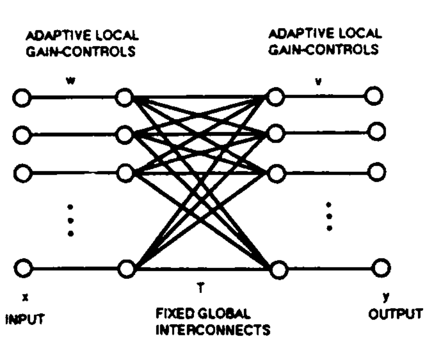


Fig.1 TAG neural network architecture

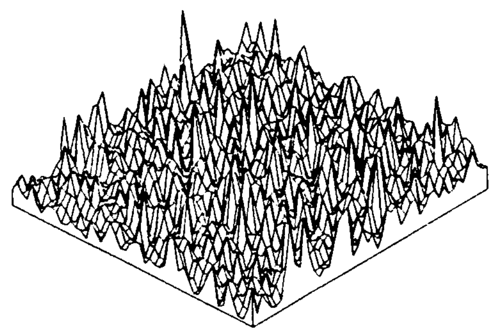


Fig.2 Diffraction indensity pattern of a pencil beam

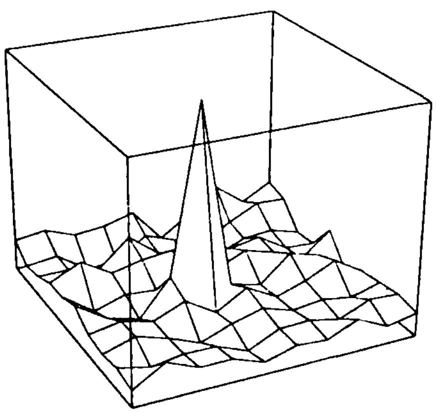


Fig.3 Correlation between diffracted patterns

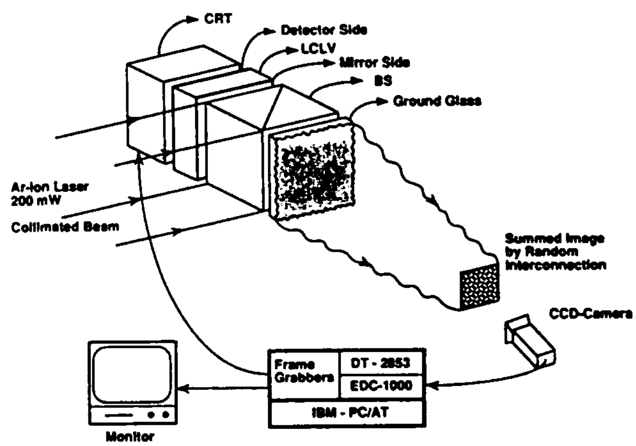


Fig.4 Setup for electro-optic implementation of TAG model

We found cheap ground glass could provide high-density random interconnections with high diffraction efficiency. In Fig.2 diffracted light intensities of a pencil laser beam from the ground glass is plotted to show irregular intensity pattern. To make sure random interconnections, in Fig. 3, correlations between intensity patterns diffracted from different ground glass points are shown as functions of distances between points. The high peak at the center represents auto-correlation, and all other cross-correlations are much smaller than the auto-correlation. Both figures clearly show random characteristics of the diffractive interconnections with ground glass.

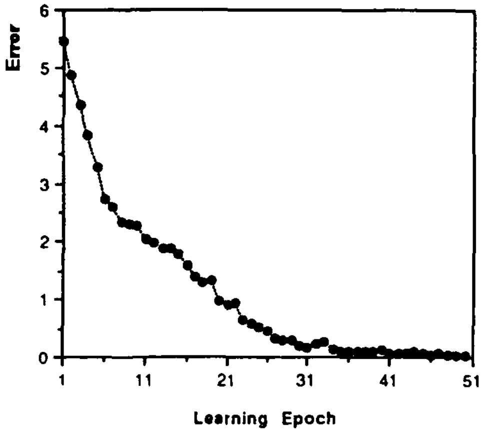
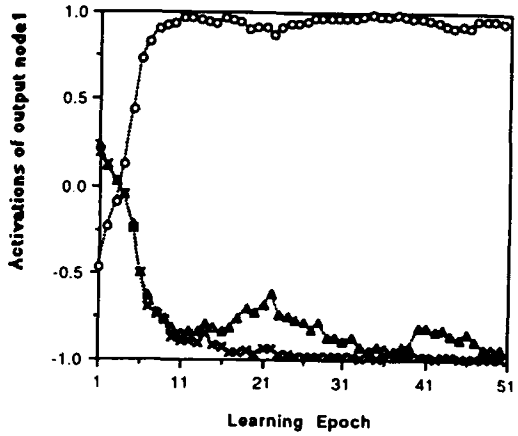
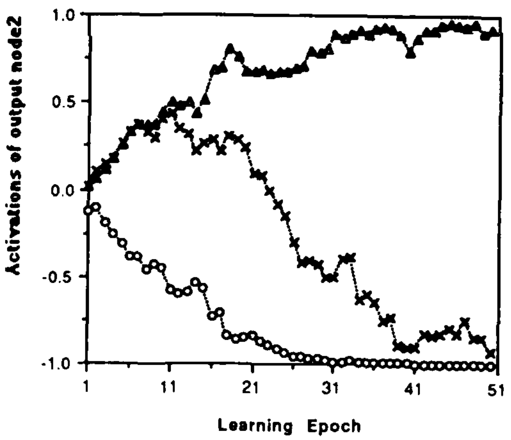


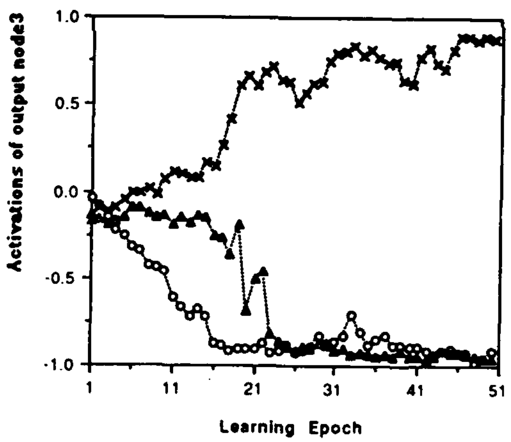
Fig.5 Error vs. learning epoch



(a)



(b)



(c)

Fig.6 Activations of the output neurons during learning. (a) the 1st neuron; (b) the 2nd neuron; (c) the 3rd neuron

4. Optical Implementation

Fig. 4 shows schematic illustration of electro-optical implementation of the single-layer TAG model. For simplicity only the forward signal path was actually implemented by optics here. The local gain-controls w_{kl} are combined with input x_{kl}, and displayed at a cathode ray tube (CRT) with grey levels. A Liquid Crystal Light Valve (LCLV), with its detector side facing the CRT and a uniform collimated laser beam illuminating its mirror side, generates a 2-dimensional pattern corresponding to $w_{kl}x_{kl}$. The ground glass diffracts with random diffraction patterns, and the diffracted lights are collected by a carge coupled device (CCD) camera. The output gain-controls v_{ij} are implemented by a personal computer (PC). The error back-propagation and gain adjustments are also implemented by the PC. However, due to its scalar nature of gradient calculations in Eqs. (3) and (4), the PC may be easily substituted by 2-dimensional arrays of smart pixels consisting of photodiodes and electronic circuits. It is worthy noting that neither lens nor hologram is used in this module. The simplicity of this architecture with ground glass is a big advantage over existing architectures using volume hologram, lenslet array with SLMs, or page-oriented holograms. For multi-layer TAG one need cascade this module.

In the experiment we chose a simple optical system to classify three 5×6 number patterns, i.e. "0", "1", and "2". Therefore only 5×6 input neurons and 3 output neurons are required. In practice, to simulate bipolar synapses, difference between two output neuron values are used.

In Fig.5 total output error, defined in Eq.(2), is shown as a function of learning epoch. Activations of 2 output neurons are also plotted as functions of learning epoch in Fig.6. Both figures show adaptive learning capability of the TAG model with ground glass.

5. Conclusion

In this paper we have demonstrated feasibility of large-scale electro-optic implementation with TAG neural network architecture and ground glass. Although more investigations are due for practical and optimal selection of ground glass particle size, this optical architecture is much more simpler than existing architectures. Provided 2-dimensional arrays of photodiodes and electronic circuits were available, this TAG with ground glass may come up with compact solid electo-optic artifical neural networks devices for classifying large amount of high resolution patterns.

Acknowledgement: This research was supported by Korea Science and Engineering Foundation.

References

[1] N.H. Farhat, D. Psaltis, A. Prata, and E. Paek, "Optical implementation of Hopfield model," *Appl. Opt.* **24**, 1469-1475, 1985.

[2] H.J. Caulfield, J. Kinser, and S.K. Rogers, "Optical neural networks," *Proc. IEEE* **77**, 1573-1583, 1989.

[3] W.T. Cathey, S. Ishihara, S.Y. Lee, and J. Chrostowski, "Optical information processing systems," *Japanese IEICE Trans. Fundamentals* **E75-A**, 28-37, 1992.

[4] J.S. Jang, S.W. Jung, S.Y. Lee, and S.Y. Shin, "Optical Implementation of the Hopfield Model for Two-Dimensional Associative Memory," *Optics Letters* **13**, 248-250, 1988.

[5] D. Brady, X.G. Gu, and D. Psaltis, "Photorefractive Crystals in Optical Neural Computers," *SPIE Proc. 882 Neural Network Models for Optical Computing*, 132-136, 1988.

[6] J.S. Jang, S.Y. Shin, and S.Y. Lee, "Programmable Quadratic Associative Memory Using Holographic Lenslet Arrays," *Optics Letters* **14**, 838-840, 1989.

[7] H.J. Lee, S.Y. Lee, and S.Y. Shin, "TAG: A neural Network Model for Large-Scale Optical Implementation," *Neural Computation* **3**, 135-143, 1991.

[8] S.Y. Lee, H.J. Lee, and S.Y. Shin, "Modified TAG neural network for large-scale optical implementation," *SPIE Proc. 1812 Optical Computing and Neural Networks*, 1992.

Analog Optical Processing

OWD 4:00pm-6:10pm
Mesquite Room

Soo Young Lee, *Presider*
Korea Advanced Institute of Science & Technology, Korea

Optical Processing in Photorefractive Crystals

John H. Hong
Rockwell International Science Center
MS A25A, 1049 Camino dos Rios
Thousand Oaks, CA 91360-2362

We describe the use of photorefractive crystals in several rf signal processing applications where the unique capability of dynamic holography operating in conjunction with acousto-optic devices provides effective processing systems.

Photorefractive Phased-Array-Radar Processor Dynamics

Robert T. Weverka, Anthony W. Sarto, and Kelvin Wagner
NSF ERC for Optoelectronic Computing Systems
University of Colorado, Boulder

Introduction

Large adaptive, two dimensional phased-array radar antennas can consist of thousands of antenna elements, have GHz bandwidths, and must be able to steer and adapt the antenna beam rapidly in a dynamic signal environment. This represents an extremely demanding signal processing task. For example, broad band processing of a 10,000 element array, with 100 time delay samples each, at GHz rates can require 10^{15} multiplies per second. The calculation and updating of the adaptive, complex weights is well suited to three-dimensional optical processing techniques where two of these dimensions are used to represent the two dimensional topology of the radar array, and the temporal variation of the signals constitutes the third dimension. We are developing a class of optical phased-array radar processors which use three-dimensional volume holograms in photorefractive crystals to time-integrate the adaptive weights to perform beam-steering and jammer-cancellation signal processing tasks[1,2]. These processors use relatively simple components requiring only a single photorefractive crystal, two single-channel high-speed detectors, and one or two single channel acoustooptic Bragg cells. The bandwidth capabilities of these components approach a GHz, allowing the processing of wide-band phased-array signals. The required number of processor components is independent of the number of elements in the phased-array. Thus in contrast to traditional electronic or acousto-optic approaches[3,4], the hardware complexity of the processor does not scale in proportion to the size of the array.

In this paper we describe an adaptive, null-steering phased-array optical processor that computes the angles of arrival of multiple interfering radar jammers and extinguishes them. The only information we assume *a priori* is that the desired signal is broad band and the jammers are narrowband. The jammer suppression dynamics have been theoretically modeled and experimentally verified.

The phased-array-radar jammer-nulling system is depicted in figure 1. The radar signals are coherently converted to the optical domain and imaged onto a photorefractive crystal (PRC). The signal which passes undiffracted through the crystal is heterodyne detected, delayed and passed to the output of the processor. A portion of the output is fed back to the acoustooptic Bragg cell to write gratings in the PRC. The gratings in the PRC diffract some of the signal incident from the phased array onto a second detector. This signal is heterodyne detected and subtracted from the undiffracted signal. A strong narrow-band jammer signal incident on the system will be correlated with the delayed version of itself in the feedback Bragg cell and hence produce a stationary interference pattern on the photorefractive crystal. The time integrating photorefractive crystal builds up gratings proportional to this stationary interference pattern. As the grating builds up, a portion of the incident light is diffracted off of the grating, producing a jammer estimate which is electronically subtracted from the undiffracted signal. This reduces the jammer content in the feedback Bragg cell making the grating build up more slowly. The grating continues to build up, reducing the jammer content in the output, until at steady state the residual jammer content in the output has been reduced by the reciprocal net gain around the feedback loop. In contrast, wideband signals are de-correlated after the feedback delay so do not write gratings, and hence are not nulled by the system. In addition, angularly resolvable inputs are Bragg mismatched so are not affected by the gratings due to other inputs.

Processor Feedback Dynamics.

A study of the feedback dynamics of the processor yields useful predictions regarding both the temporal behavior and steady-state values of jammer suppression. Such an analysis provides information on how the dynamic convergence depends on the relative phases and amplitudes of the optical signals, as well as on the electrical gains and delays. The analysis establishes the expected depth of nulls, and reveals the range of frequencies, determined by the resulting phase relationships, that result in effective jammer nulling. The analysis includes the effects of phase relationships between optical input and reference signals, and phase shifts due to electrical and acoustic delays. This is done by analyzing the dynamics of the complex variable G, the grating strength in the PRC. A subsequent transformation of the rate equations to equations describing the normalized jammer excision ratio, E, which is proportional to the residual jammer content in the system output, provides a transparent description of both the dynamic and steady state behavior.

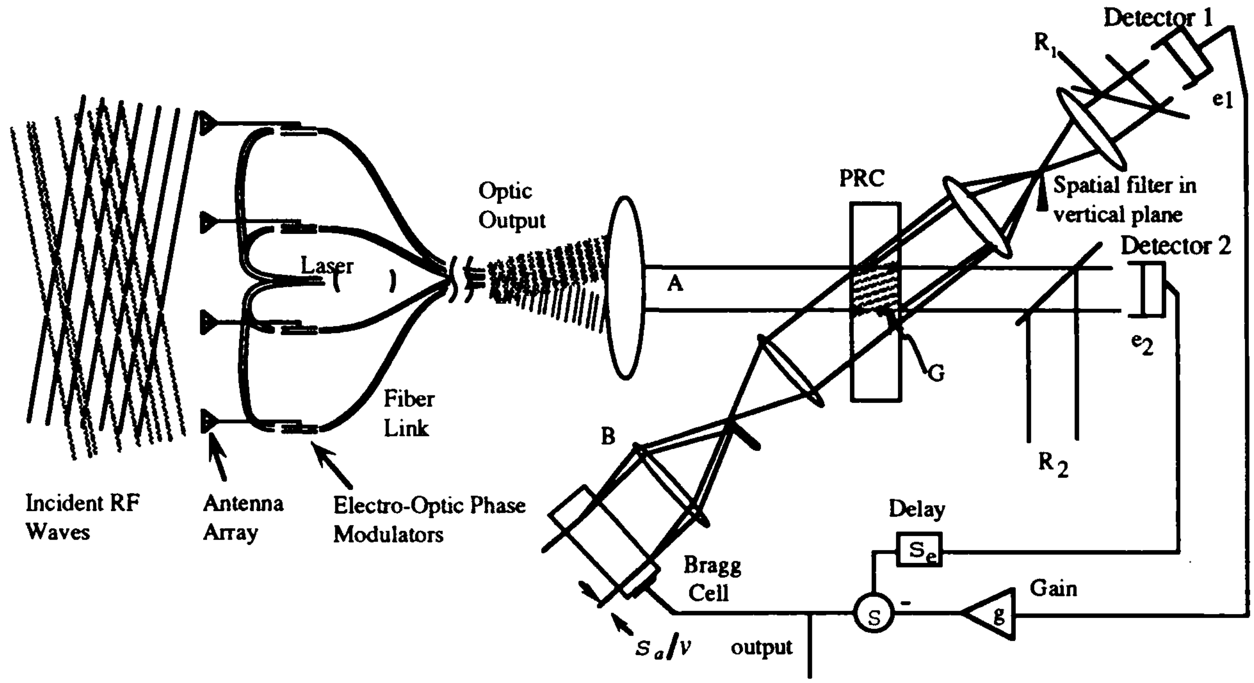


Figure 1. Photorefractive phased array radar processor and fiber remoted input.

A schematic of the processor feedback system is shown in figure 1. The input to the processor uses point electrooptic phase modulators, one for each element of the phased-array, to convert the RF signals into the optical domain[5]. The light from the phase modulators is transmitted to the processor through a fiber optic feed network and imaged onto the photorefractive crystal. The signal $A = A_0 e^{i\omega t}$ in the figure represents the optically modulated phased array radar signal due to a plane wave jammer at frequency $\omega/2\pi$. The component of A which is diffracted by the Bragg matched grating, G, in the PRC is detected at detector 1 and the transmitted component by detector 2 with heterodyne references $R_1 e^{i\phi_{R1}}$ and $R_2 e^{i\phi_{R2}}$ respectively, resulting in signals e_1 and e_2 given by

$$\begin{aligned} e_1(t) &= \left|GA_0 e^{i\omega t} + R_1 e^{i\phi_{R1}}\right|^2 \Re \\ e_2(t) &= \left|A_0 e^{i\omega t} + R_2 e^{i\phi_{R2}}\right|^2 \Re \end{aligned} \qquad (1)$$

where $\Re$ is the detector responsivity.

The signal e_1 is amplified by the factor g and subtracted from a delayed version of e_2. This is the RF input into the Bragg cell. The optical feedback signal, now with an additional acoustic delay σ_a, is diffracted into the +1 order of the Bragg cell, and is given by

$$B(t) = A_0 C\eta\Re R_2 e^{-is} e^{i\omega(t-x/v)}\left(1 - g(R_1/R_2)Ge^{ip}\right) \qquad (2)$$

where h is the small-signal acoustooptic diffraction efficiency, $s \equiv \phi_{R2} - \phi_C - \omega\sigma_e - \omega\sigma_a$ is the phase at detector 1, $p \equiv \phi_{R2} - \phi_{R1} - \omega\sigma_e$ is the phase at detector 2, v is the acoustic velocity and C is the input illumination amplitude of the Bragg cell. The PRC responds to intensity, which at the crystal is given by $|A+B|^2$. The spatially periodic portion of the optical intensity writes the photorefractive grating. The grating decays in proportion to the existing grating and the average intensity. Thus, the temporal response of the complex grating G is given by

$$\frac{\partial}{\partial t}G = -\alpha\left(|A|^2 + |B|^2\right)G + \beta\left(A^* B\right) \qquad (3)$$

where a and b are the erasure and writing sensitivities of the PRC. Making the substitutions $a = 1/\left(I_C I_{R2}\eta^2\Re^2\right)$, $q \equiv s - p = \phi_C - \phi_{R1} + \omega\sigma_a$, $b = gR_1\beta/\left(\alpha C I_{R2}\eta\Re\right)$, $E \equiv 1 - gGe^{ip} R_1/R_2$, $t' = t\left(\alpha I_A I_C I_{R2}\eta^2\Re^2\right)$ equation (3) becomes

$$\frac{\partial}{\partial t'}E = \left(a + |E|^2\right)(1-E) - be^{iq}E \qquad (4)$$

where the dimensionless (in general complex) variable E, is proportional to the normalized residual jammer content in the feedback signal. Note that at the onset of a jammer, E is equal to 1, and proper operation of the processor drives E to a small value. We have obtained numeric results for equation (4), and analytic solutions for a linearized version of (4). Near convergence, i.e. for small E, the linearized version yields an analytic expression for the steady state jammer suppression, E_{ss}, and gives a decay time constant.

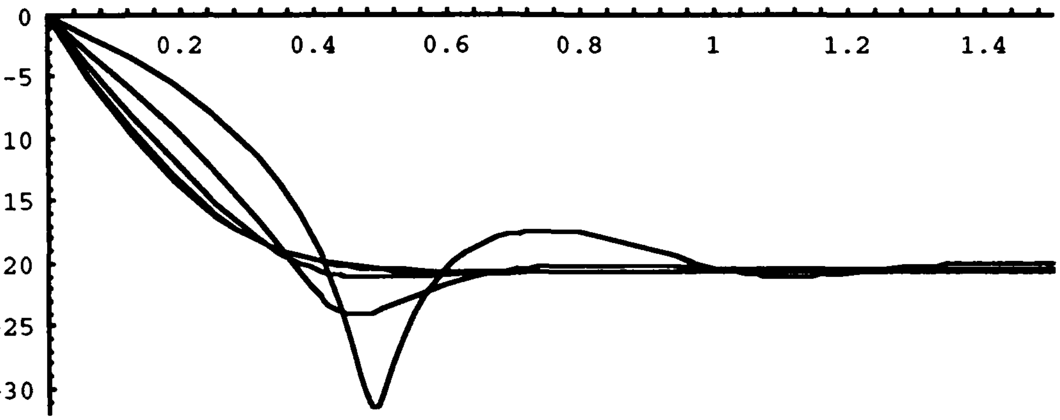


Figure 2. Plot of numeric solution of equation (4), jammer suppression vs. time for q = 0p, .1p, .2p, .3p, .4p.

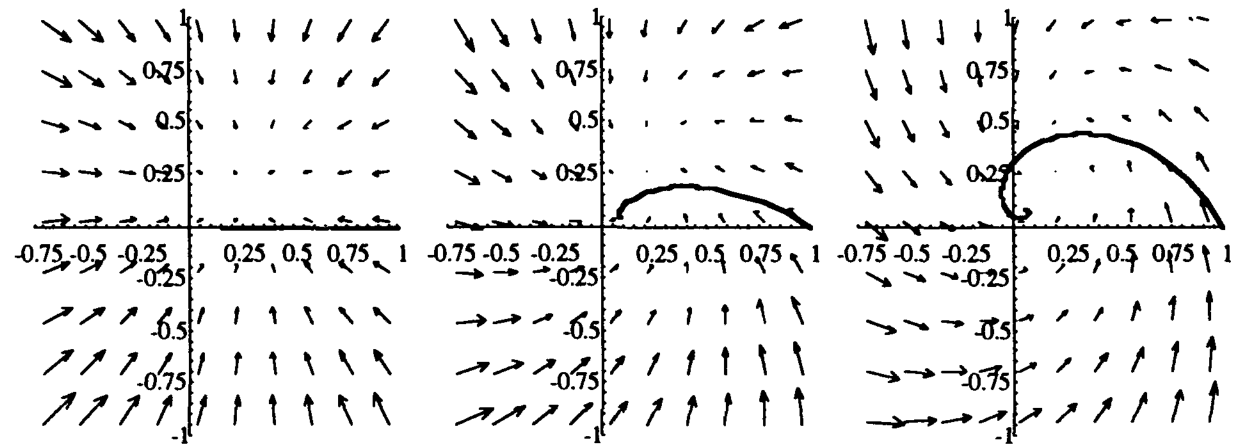


Figure 3. Plot of numeric solution of equation (4) showing Re[E] versus Im[E] parameterized in time,superimposed over vector field plot of the velocity from equation (4), for q = 0, p/6, p/3.

$$E_{ss} = a/b\left(1 + 2\,a/b\cos(q)\right)^{1/2} \approx a/b \qquad (5)$$
$$\tau^{-1} = a + b\cos(q)$$

Equation (5) is valid for $-\pi/2 \le q \le \pi/2$ and small a/b (large feedback gain). The feed back phase q has a linear frequency dependence due to the turn-on acoustic transit delay in the Bragg cell, σ_a. The finite range q for which we have convergence limits the system bandwidth, B, to half this reciprocal delay, $B \le 1/(2\sigma_a)$.

Figure 2 plots the numeric solution of the full nonlinear expression of $E(t')$ given by (4) for a = 1, b = 10 and parameterized in q from 0 to p/2 in steps of p/10. The first four solutions converge rapidly, while the convergence at q =0.4p is slightly slowed.

Additional insight into the temporal dynamics of $E(t')$ is obtained by plotting the complex $E(t')$ parameterized in t', superimposed on a vector field plot of the velocity from (4) for increasing values of q as shown in figure 3. The convergent behavior for q up to p/2 is evident. Note that for q = 0, E is purely real, thus, E travels only along the real axis.

Thus, it is clear that the value of q, i.e. the relative phase relationship between j_c, j_{RI} and the phase due to the acoustic delay ws_a, has a significant effect on both the temporal response and the absolute value of jammer suppression. For effective jammer suppression, a q value in the range of -p/2 < q < p/2 is necessary, however as shown in the small a/b limit the jammer suppression is independent of q. In addition the frequency dependent phase term ws_a requires that s_a be reduced for broad band operation.

Experimental Results

We have constructed the system described above using an optical simulator for the phased-array radar input. The simulator for a far field jammer consists of an acoustooptic modulator at the back focus of a lens. Light from the modulator is collimated and sampled by a Ronchi ruling to simulate the fiber apertures in the fiber remoting system. The system uses two large-area high-speed detectors, a large time-bandwidth Bragg cell

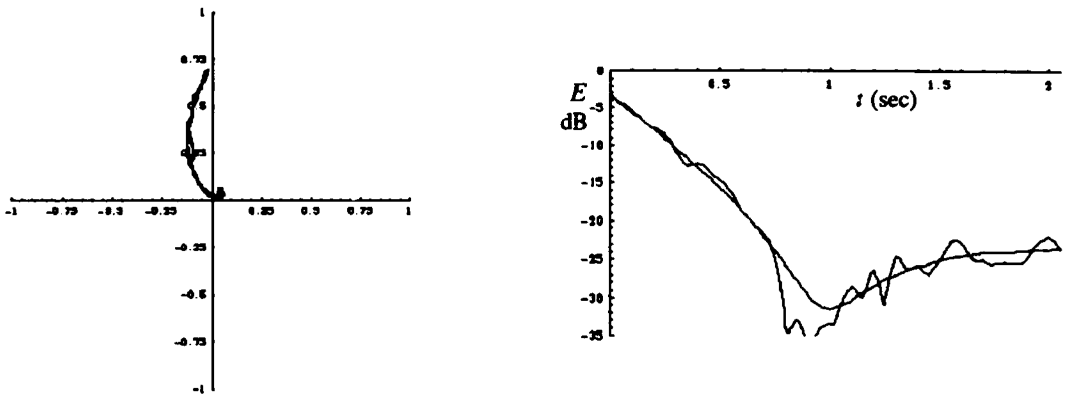


Figure 4. a)Complex excision and b) Magnitude of the excision versus time, experiment and fit

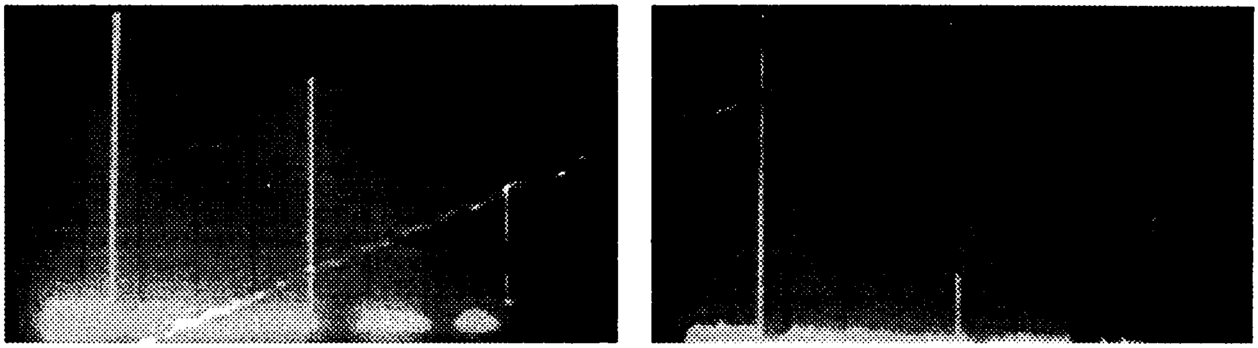


Figure 5. Spectrum analyzer trace with feedback off (left), and on (right), showing 35 dB of jammer suppression. The scale is 10 MHz per division horizontal, 10 dB per division vertical.

in the feedback loop, an electronic phase shifter in the feed back path to set the operating phase, and a $BaTiO_3$ photorefractive crystal as the volume holographic storage media. The system operates between 50 MHz and 100 MHz with an instantaneous bandwidth of 1 MHz limited by the acoustic delay from the transducer to the optical beam.

The output of the processor is mixed down electrically with both in-phase and quadrature reference signal taken from the jammer input. This allows the measurement of the complex excision, E. Figure 4 shows the experimental measurements of E from turn-on of the jammer. Figure 4a shows the motion of E in the complex plane corresponding to the theoretical plot of Figure 3, and Figure 4b shows the decay of the magnitude of E. The processor output is also fed to a spectrum analyzer where the magnitude of the jammer excision is measured. Figure 5 shows -35 decibels of excision obtained using large feedback gain.

Conclusion

We have derived expressions for the dynamics and steady state suppression of the phased-array-radar processor for the case of a single strong jammer. The system has a high suppression ratio, and a fast convergence for a wide range of the feedback phase, and hence a large bandwidth. We have presented experimental verification of the system behavior showing agreement with the theoretical model and showing jammer suppression as high as -35 dB.

Acknowledgments

This work was supported by Rome Labs contract number F30602-92-C-0011.

References

[1] Weverka, R.T. and Wagner, K., Staring phased-array radar using photorefractive crystals , in *proc. SPIE vol. 1564-63* , San Diego, July 1991.

[2] Weverka, R.T. and Wagner, K. Adaptive phased-array radar processing using photorefractive crystals. In *proc SPIE vol. 1217-33,* Jan. 1990, pp. 173-182.

[3] Voskresenskii, D.I., Grinev, A.I., and Voronin, E.N., *Electrooptical Arrays.* Springer-Verlag, 1989.

[4] Psaltis, D. and Hong, J., Adaptive acoustooptic processor, *Proceedings of the SPIE 519*(936), 62.

[5] Guggenmos, J., Johnson, R., and Seery, B., Optical distribution manifolds for phased-array antennas, in *proc. SPIE vol. 886,* 1988.30602-

Performance Evaluation of an Acousto-Optic Wide Band Correlator System

R.D. Griffin and J.N. Lee
U.S. Naval Research Laboratory
Optical Sciences Division, Code 5623
Washington, D.C. 20375-5320
202-767-3100, 202-767-6370 (FAX)

We have designed and built an acousto-optic (AO) correlator system that is used as a matched filter processor integrated into a digital signal processing system, and have analyzed parametrically the efficiency of correlating a continuous time series waveform. Acousto-optic technology has the potential to provide a 100- to 1000-fold improvement in matched-filter processing power over the fastest digital systems with significantly lower volume and power requirements. The optical engine in our system has the potential to provide the equivalent of 8-10 GFLOPS of processing power. With its current electronic interface package, less than one percent of the optical potential is harnessed. Nevertheless, it performed 20 to 70 times faster than a VAX 6410 using a vector processor and an optimized FFT correlation routine. Thus, it has the potential to run several thousand times faster than the VAX. The optical system and electronic interface occupy 1.5 ft^3 and consume approximately 200 W of power. This yields a potential system figure of merit of 27-33 GFLOPS/ft^3/kW. Two-dimensional optical techniques have the potential to enhance the computation rate by a factor of 25 to 100, and improved electronics and packaging can further improve the figure of merit. Integrating such a powerful computational engine into a digital signal processing system is a significant challenge due primarily to the large input and output data bandwidths, but also to the necessary conversions between the analog and digital domains.

The optical portion of our AO correlator system is controlled by a set of interface electronics. The electronics package contains buffer memory, digital-to-analog converters (DAC's) and amplifiers to generate the waveforms, a system controller to set up and supervise the hardware and communicate with the host, and an analog-to-digital converter to digitize the correlation functions. The system is described in more detail in Ref. 1.

The basic optical correlation operation is shown in Fig. 1 where two reference waveforms are shown propagating to the right and two receive waveforms propagate to the left. In our system, these waveforms propagate through acousto-optic cells that are illuminated by an optical beam. When the reference waveform overlaps the receive waveform inside the optical aperture (represented by the duration T_{OA}), the light diffracted by the signals in the AO cells effects a multiplication of the reference and receive waveforms. A lens then integrates the multiplied signals onto a photodetector where the correlation signal is detected. For more details, see Ref. 2.

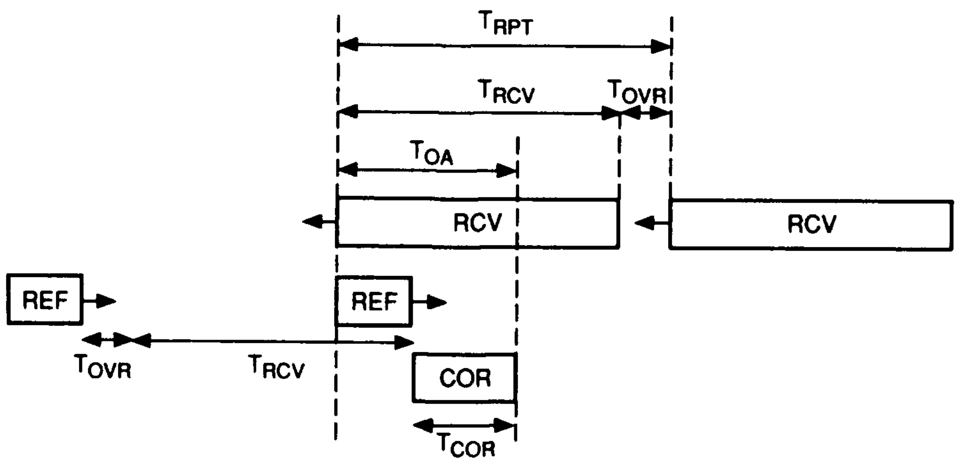


Fig 1.Optical correlation timing sequence.

We can use Fig. 1 to derive a measure of how efficiently we are using the optical processor. Since the correlation signal is valid only when the reference waveform overlaps the receive waveform inside the optical aperture, and since the waveforms are *counter* propagating, the duration of the correlation signal is $T_{COR} = \frac{1}{2}(T_{RCV} - T_{REF})$, as shown in the figure. The time between correlations, i.e., the correlation period, is indicated by $T_{RPT} = T_{RCV} + T_{OVR}$, where T_{OVR} represents any overhead time. One can also see from the figure that, for a single correlation segment, the longest possible receive waveform satisfies $T_{RCV}^{max} = 2T_{OA} - T_{REF}$. Our system uses digital waveforms that are converted to analog values at a rate f_{CLK} = 80 MHz, and has an optical aperture T_{OA} = 44 μs. Because of the counter propagation, the correlation signal must be digitized at a rate $2f_{CLK}$. We define the average computation rate as the ratio of the number of correlation samples produced to the correlation period,

$$f_{avg} = \frac{2f_{CLK}T_{COR}}{T_{RPT}} = f_{CLK}\frac{T_{RCV} - T_{REF}}{T_{RCV} + T_{OVR}}. \tag{1}$$

For a given value of T_{REF}, this has its peak value when the overhead is zero and the longest receive waveform is used,

$$f_{avg}^{max} = 2f_{CLK}\frac{T_{OA} - T_{REF}}{2T_{OA} - T_{REF}}. \tag{2}$$

We define the efficiency η as the ratio of the average computation rate to the peak computation rate,

$$\eta = \frac{f_{avg}}{f_{avg}^{max}}. \tag{3}$$

The efficiency parameter of Eq. (3) provides a measure of the effect of overhead on processor performance. The overhead time depends on the hardware setup time, on the time required to transfer data to or from the host during a correlation period, and on the amount of code that must be executed by the system controller between each correlation period. The correlator supports three subroutines that perform correlations. The first is a diagnostic routine that continously correlates a test waveform against a zero-padded copy of the waveform, i.e., it autocorrelates the test waveform. The second subroutine cross-correlates a receive waveform against a user-specified number of references. It digitizes and stores the correlation functions, and, if requested, it will return the digitized correlation functions to the host. The third subroutine compares correlation functions with threshold functions and reports any detections (threshold crossings) to the host.

The diagnostic correlation subroutine contains a small amount of code that performs a simple computation, sets up the various hardware registers and counters, initiates the correlation, and then waits for the end of correlation signal before beginning again. The subroutine correlates a 2048 sample waveform against a simulated receive waveform consisting of 2048 zeroes, concatenated with a copy of the waveform, concatenated with enough zeroes to make the receive waveform as long as possible. Thus, the receive waveform effectively contains 4992 samples. Under these conditions, the observed correlation period was T_{RPT} = 189 μsec, which implies an overhead time of 127 μsec. The resulting values for Eqs. (1), (2), and (3) are shown in Table 1. The efficiency of 33 percent indicates the necessity for asynchronous software, i.e., software that executes during the correlation to set up the registers. Even with asynchronous operation, the software must be executed in half the time if the overhead time is to approach zero. In practical terms, this means the code must be written more efficiently and the digital processor in the system controller must be replaced with a faster one.

Measurements made on the other two correlation subroutines indicate that they are operating at a bit less than one percent efficiency. This efficiency refers to the time between consecutive correlations; it does not include the time required to transfer data to the host. If it did, the efficiency would drop by factors of approximately 2 to 10, depending upon whether the host was a VAX 6000 or an IBM PC.

Table 1. Performance of the continuous correlation (diagnostic) subroutine.

Conditions:	$T_{REF} = 25.6\ \mu$sec $T_{RCV} = 62.4\ \mu$sec $T_{RPT} = 189\ \mu$sec	(2048 samples) (4992 samples)
Avg. no. correlation samples/sec, Eq. (1)	f_{avg}	15.6×10^6 smp/s
Max. avg. no. correlation samples/sec, Eq. (2)	f_{avg}^{max}	47.2×10^6 smp/s
Efficiency, Eq. (3)	η	0.33

We can also define further measures of performance. The raw computation rate is the rate at which multiplications are done to produce one correlation function sample. Since the number of multiplications done to produce one correlation function sample is the number of replica samples N_{REF}, the raw rate is given by

$$f_{raw} = \frac{N_{REF}}{1/(2f_{CLK})} = 2f_{CLK}^2 T_{REF} \tag{4}$$

We have shown elsewhere [1] that the required digital processing power (floating point operations per second) to equal the optical correlator operating in its most efficient mode is given by

$$f_{\mathrm{FLOPS}}^{effic} = 4f_{CLK}\frac{(1-\tau_{REF})}{(2-\tau_{REF})}\frac{N_{opt}\left(\log_2 N_{opt}+\frac{1}{2}\right)}{\left(N_{opt}-N_{REF}\right)}, \tag{5}$$

where $\tau_{REF} = T_{REF}/T_{OA}$ and N_{opt} is the optimum size [3] for the FFT algorithm used to perform the correlation. The results of Eqs. (2), (4), and (5) are shown for a few values of reference waveform duration in Table 2. The raw computation rate is typically several hundred billion multiplications per second. This yields a maximum average computation rate of fifty million correlation samples per second, which would require at least 2 GFLOPS in a digital processor. A practical rule of thumb for a digital system is that this number must be increased by a factor of four or five to overcome the inefficiencies that inevitably appear in a digital system.

When connected to a VAX computer, the user has the option of linking the application code to a library of emulation subroutines. These are subroutines that emulate the functions of the optical correlator in software. Two versions of the emulator library are available: a version that uses the standard VAX scalar processor, and a version that uses the VAX vector processor (if it is available on the user's VAX). Since these subroutines execute directly on the VAX, they have no I/O overhead associated with transferring data over the parallel interface to and from the optical correlator (although there may be I/O overhead associated with virtual memory). Since the emulator subroutines are functionally identical to the optical correlator subroutines, the user can use them to make various comparisons between the digital (emulator) correlator and the optical correlator.

In one set of tests, we compared the performance of three equivalent correlators: the optical correlator connected to the VAX, an FFT correlator using the VAX vector processor, and an FFT correlator using the normal VAX scalar processor. The VAX was a VAXVector 6410, that is, a VAX 6410 with an FV64A vector processor installed. Data communication with the correlator was via a 16-bit DRB32W parallel I/O port. The same code ran in each test, except that subroutine calls to the optical correlator were replaced by calls to emulation subroutines for the other two correlators. The same data sets were run on all three correlators. The results are shown in Tables 3 and 4. These tables list the cpu time charged to the digitize subroutine for each correlator and compare these times in two ways: Table 3 compares the cpu time for the vector and scalar correlators with the optical correlator; Table 4 compares the times as a function of the number of correlations performed per subroutine call. The uncertainties in Table 3 indicate one standard deviation: they are due primarily to activity from other processes on the VAX (running VMS). In Table 3, as the number of correlations increases, the time ratios approach steady state values. We can take these ratios as measures of the performance improvement available with the optical correlator.

The Ratio columns in Table 4 show that, as one might expect, the processing times for the scalar and vector correlators scale approximately linearly with the number of correlations. This is not the case with the optical correlator, however. Since the actual correlation time is negligible, the increase in time is due almost entirely to data transfer time and the execution time for the control software.

This work was sponsored by the Office of Naval Technology and the Office of Naval Research.

Table 2. Computation rates ($f_{CLK} = 80$ MHz).

τ_{REF}	f_{raw} (10^9 mul/s)	f_{avg}^{max} (10^6 smp/s)	f_{FLOPS}^{effic} † (10^9 FLOPS)
0	0	80.0	2.47
0.1	56.3	75.8	2.34
0.25	141	68.6	2.12
0.33	186	64.2	1.98
0.5	282	53.3	1.65
0.67	377	39.7	1.22
0.75	422	32.0	0.987
1.0	563	0	0

† $N_{REF} = 1024$ and $N_{opt} = 8192$.

Table 3. Comparison of optical, vector, and scalar correlators.

No. Correls.	t_{Opt} (ms)	t_{Vector} (ms)	t_{Vector} / t_{Opt}	t_{Scalar} (ms)	t_{Scalar} / t_{Opt}
1	18±9	320±23	18	991±18	55
2	19±8	714±40	38	2021±39	106
4	26±9	1415±43	54	4232±119	163
6	32±10	2136±56	67	6274±185	196
8	41±12	2767±60	67	8294±198	202

Table 4. How the optical, vector, and scalar correlators scale with number of correlations.

No. Correls.	t_{Opt} (ms)	Ratio	t_{Vector} (ms)	Ratio	t_{Scalar} (ms)	Ratio
1	18	1	320	1	991	1
2	19	1.06	714	2.2	2021	2.04
4	26	1.4	1415	4.4	4232	4.3
6	32	1.8	2136	6.7	6274	6.3
8	41	2.3	2767	8.6	8294	8.4

References

1. R.D. Griffin and J.N. Lee, chapter in *Acousto-Optic Signal Processing, Theory and Implementation: 2nd Edition* (Marcel Dekker, to be published).
2 A. VanderLugt, *Optical Signal Processing* (John Wiley & Sons, Inc., New York, 1992), Ch. 13.
3 H.D. Helms, "Fast Fourier transform method of computing difference equations and simulating filters," IEEE Trans. Audio and Electroacoustics **AU-18**, 439-441 (Dec. 1970).

Optical Matrix Multiplication using Grating Degeneracy in Photorefractive Media

Claire Gu
Department of Electrical and Computer Engineering
The Pennsylvania State University
University Park, PA 16802
(814)863-4256

Scott Campbell and Pochi Yeh*
Department of Electrical and Computer Engineering
University of California at Santa Barbara
Santa Barbara, CA 93106
(805)893-3981

Matrix-matrix multiplication is an important operation in many computational and processing applications including correlation, convolution, Fourier transform of temporal signals and 2-dimensional images. In addition, a large number of signal and image processing algorithms can be expressed in terms of matrix operations. Direct matrix-matrix multiplication is often avoided in electronic computers because it is an $O(N^3)$ (where $N \times N$ is the number of elements in each matrix) operation which requires a long computation time for serial machines. Optical computing offers the advantages of parallelism and large capacity. Such capabilities have been successfully demonstrated in parallel vector-matrix multiplication[1]. Although matrix-matrix multiplication can be performed as an extension of vector-matrix multiplication[1] with color multiplexing or time multiplexing, these schemes are complicated by dispersion or time delay. Recently, nonlinear optical techniques have been employed in the parallel matrix-matrix multiplication[2,3]. These techniques require complicated alignment and suffer from severe energy loss. In this paper, we propose and demonstrate a new method which utilizes grating degeneracy[4] in photorefractive media in conjunction with an incoherent laser array to implement parallel optical matrix-matrix multiplication. Specifically, multiplications are implemented by photo-induced index gratings whose amplitudes are determined by the interference between coherent beams, while summations are implemented by grating degeneracy. Such a matrix-matrix multiplier is capable of handling large matrices such as those with 1000×1000 elements.

Fig. 1 shows the schematic diagram which describes the principle of operation of the matrix-matrix multiplication. Both matrix A ($N \times N$) and matrix B ($N \times N$) are placed at the front focal plane of lens L_1. At the rear focal plane of lens L_1, a volume holographic medium such as a photorefractive crystal is inserted to record the multiplication of the two matrices. The recorded information is read out by a set of reading beams which consists of N diagonally aligned point sources placed at the front focal plane of lens L_2. Notice that the recording beams are propagating from left to right while the reading beams are propagating from right to left. The holographic medium is also located at the rear focal plane of lens L_2. The diffracted readout beams are directed by a beamsplitter to the output plane which is located at the focal plane of lens L_1.

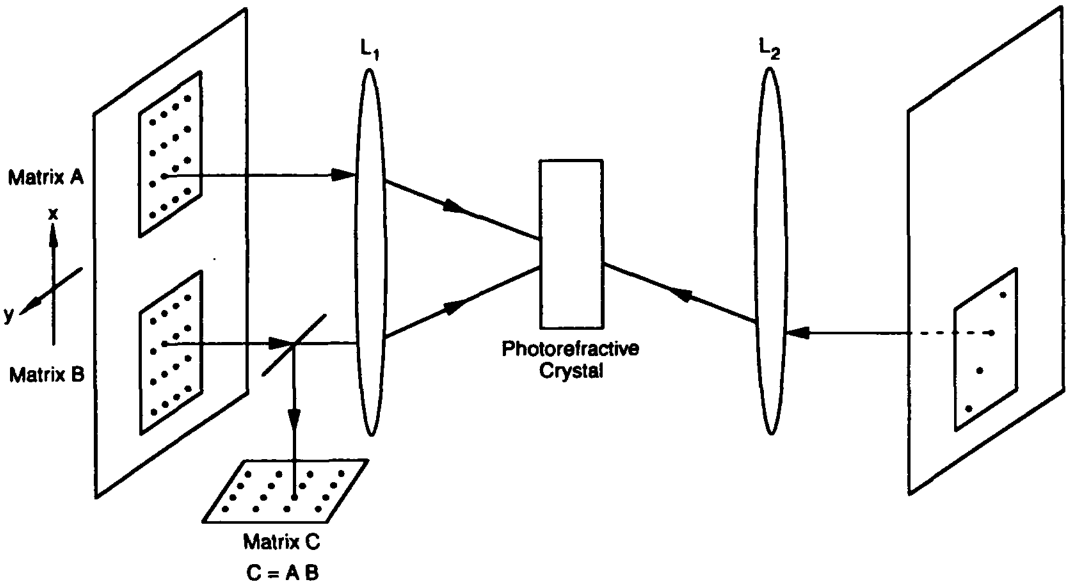


*Fig.*1 Schematic diagram for matrix-matrix multiplication.

The matrix elements are arranged in the following order, as shown in Fig. 2(a). In the region of matrix A, each line along the y-direction represents a row of matrix A. In the region of matrix B, each line along the y-direction represents a column of matrix B. To realize matrix-matrix multiplication, the illumination of the matrices A and B is chosen so that all pixels within each line along the x-direction are mutually coherent while pixels with different y values at the input plane are mutually incoherent. During the recording, gratings are formed according to $A_{ij}B^*_{jk}$. There are N^2 gratings formed within each line along the x-direction and the total number of gratings is N^3. Within these N^3 gratings, there are only N^2 different grating wave vectors, i.e., there are N degenerate gratings for each grating wave vector. The corresponding degeneracy condition is shown in Fig. 2(b) in the normal surface representation. As an example, the index grating representing the four terms in Fig. 2 is

$$\Delta n \propto \frac{1}{I_0}\mathrm{Re}[(A_{11}B^*_{14} + A_{12}B^*_{24} + A_{13}B^*_{34} + A_{14}B^*_{44})e^{-i\mathbf{K}\cdot\mathbf{r}}], \tag{1}$$

where I_0 is the total averaged intensity and $\mathbf{K}$ is the common grating wave vector. When this index grating is read out, the diffracted beam amplitude is proportional to the grating amplitude which is exactly the matrix element $C_{14} = \sum_{j=1}^{4} A_{1j}B_{j4}$. In the above derivation, we have used the weak grating approximation and neglected gratings with small grating wave vectors.

During the readout, all elements of matrix C can be read out in parallel by using N point sources that are diagonally aligned at the reading plane, as shown in Figs. 1 and 3. Each of the reading points reads out N nondegenerate gratings, giving N diffracted points in a line along the x-direction at the output plane. The $N \times N$ elements of matrix C are represented by the N^2 set of nondegenerate index gratings.

The method described above has been implemented experimentally. Experimental results are in excellent agreement with the theoretical predictions. For example, the experimental results for the

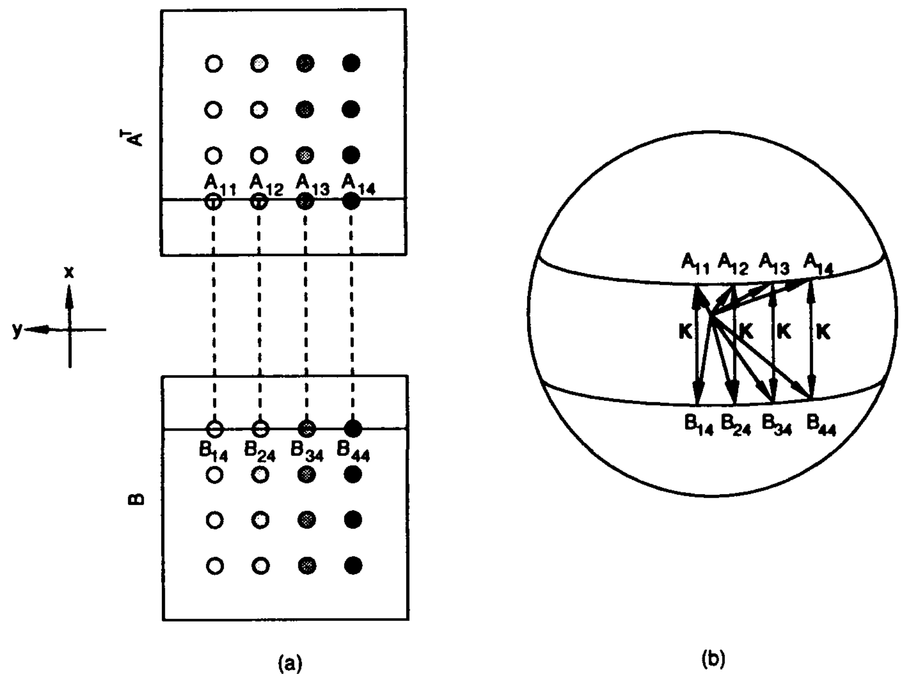


*Fig.*2 Arrangement for elements during recording. (a) At the input plane, pixels with the same shading are coherent and those with different shadings are mutually incoherent. (b) Degenerate gratings shown in the momentum space (normal surface).

following matrix-matrix multiplication

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 0 & 2 \\ 1 & 1 & 1 & 1 \\ 0 & 3 & 0 & 3 \\ 1 & 0 & 1 & 0 \end{pmatrix} \tag{2}$$

are given by

$$\begin{pmatrix} 0 & 2.0 & 0 & 2.1 \\ 0.9 & 0.9 & 1.0 & 1.0 \\ 0 & 3.0 & 0 & 3.0 \\ 1.0 & 0 & 1.0 & 0 \end{pmatrix}. \tag{3}$$

To our knowledge, this is the best experimental results ever reported on optical matrix-matrix multiplication.

The capacity of such an optical system is determined by the thickness of the crystal, wavelength of the laser, spot size of pixels, and the numerical aperture. With proper alignment, such an optical system is capable of handling large matrices, for example matrices with 1000 × 1000 elements. For a crystal of thickness $L = 0.5cm$, wavelength $\lambda = 0.5\mu m$, the angular separation between pixels (in the x-direction)

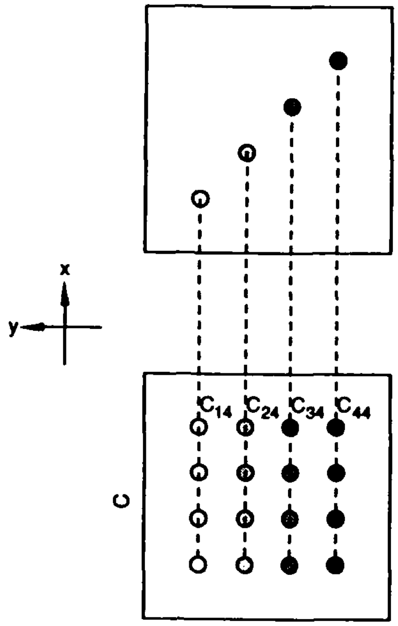


*Fig.*3 Arrangement for elements during readout. The N reading points are diagonally aligned at the reading plane. Diffraction occurs among pixels with the same shading.

can be as small as[5] $\delta\theta = \lambda/L = 10^{-4}$. Thus 1000 pixels can be placed within a numerical aperture of 0.1 which satisfies the paraxial condition. In addition, to better utilize the degeneracy condition, the y-separation between incoherent lines can be chosen as small as allowed by the diffraction limit. In this case, the pixels may have unequal horizontal and vertical separations.

In conclusion, we have proposed and demonstrated a novel method of matrix-matrix multiplication using grating degeneracy in photorefractive media in conjunction with incoherent laser source arrays[6]. The experimental results are in excellent agreement with the theoretical predictions. Such a matrix-matrix multiplier, with its large capacity and parallelism, can potentially be used in optical computing, photonic switching and optical neural networks.

Pochi Yeh and Scott Campbell acknowledge support from the Air Force Office of Scientific Research.

* Pochi Yeh is also a Principal Technical Adviser at Rockwell International Science Center.

1. J. W. Goodman, A. R. Dias, and L. M. Woody, Opt. Lett. **2**, 1 (1978).

2. J. Hong and P. Yeh, Opt. Lett. **16**, 1343(1991).

3. P. Yeh, A. E. Chiou, J. Hong, P. Beckwith, T. Chang, and M. Khoshnevisan, Opt. Eng. **28**, 328(1989).

4. H. Lee, X.-G. Gu and D. Psaltis, J. Appl. Phys. **65**, 2191(1989).

5. See, for example, H. Kogelnik, Bell Syst. Tech. J., **48**, 2909 (1969).

6. C. Gu, S. Campbell, and P. Yeh, "Matrix-matrix multiplication using grating degeneracy in photorefractive media," to appear in Opt. Lett.

Optoelectronic fuzzy logic inference system using beam scanning laser diodes

Hideo Itoh, Masanobu Watanabe, Seiji Mukai, and Hiroyoshi Yajima
Electrotechnical Laboratory, Optical Information Section
1-1-4, Umezono, Tsukuba, Ibaraki, 305 JAPAN
Email : itoh@etl.go.jp

1.Introduction

Recently fuzzy control techniques have been widely used in the control systems which are difficult to implement using conventional control methods. Various electronic approaches are reported for high speed fuzzy controllers[1]. On the other hand, application of the parallelism of optics is attractive for the improvement of the performance of fuzzy controllers. To our knowledge, however, no total system of optoelectronic fuzzy control system with high performance has been reported yet, except for a few reports on digital optoelectronic implementations of MIN or MAX operations[2].

In this paper, we propose a novel optoelectronic fuzzy reasoning method using beam scanning laser diodes[3]. The reasoning system based on the method is simpler and faster than conventional electronic systems because of its optical spatial parallelism of the steered laser beams. Antecedent, consequent, and defuzzification operations of the fuzzy inference processes are realized by scanned far-field-beam of the lasers, arithmetical sum of the beams of the lasers, and the characteristics of position sensing devices, respectively. Experimental results of each operations are shown in this paper.

2.Fuzzy reasoning methods

Inference of the fuzzy controllers are composed of grade-evaluation of antecedent fuzzy rules, modification of the consequent fuzzy functions by the evaluated grades, unification of modified functions, and the defuzzification of the unified functions. Various envelopes of membership functions and fuzzy reasoning methods can be used for system controls. In the case of high speed fuzzy controls, triangle envelope of membership functions and a reasoning method using logical products for antecedent and consequent part, logical sum for unification, and center-of-gravity for defuzzifications (MIN-MAX-gravity method[1]) are frequently used because of its easier implementations for electronic circuits. However, this architecture is not easy to implement by optical or optoelectronic fuzzy reasoning methods. Here, gaussian far-field beams of laser diodes

are used as the membership functions instead of triangle functions. As for algebraic products for antecedent and consequent part, arithmetic sum for unification is used instead of logical product and sum. Center-of-gravity method is used for defuzzification.
This product-sum-gravity method[4] is used because of its convenience for optoelectronic implementations. Figure 1 shows a configuration of the Beam scanning fuzzy inference system which deals two antecedent rules with each input-variable and a consequent rule with an output-variable. The system is composed of four beam scanning laser diodes, two photodetectors, and a position sensing photodetector. Grade evaluations of antecedent rules are realized by steered beam of the beam scanning laser diodes. Modification of membership functions of Consequent rules is realized by modulation of injected current to laser diodes of the consequent part. Unification of the result-function is given by an envelope of irradiated far-field beams on an optical beam position sensing photo-detector(PSD). Defuzzification is executed by the photodetector.

3.Experimental Results

Figure 2 (a) shows a experimental setup of a grade evaluation of antecedent rules. Ratio of injected currents to each p-electrode are changed from 0.36 to 0.64 and total injected currents are fixed at 100 mA. Distance between the beam scanning laser and photodetector array is 20 mm. Figure 2 (b) shows results of the grade calculation. Trace (A),(B), and (C) denote the cases of "Small", "Medium", and "Large",respectively.

Figure 3 (a) shows an experimental setup of result-function of results of the consequent, unification, and defuzzification operations. Injected current to a laser A is varied from 0 mA to 52 mA and current to another laser B is fixed at 50.5 mA. Separation of each laser of the laser array is 600 μm and distance between the beam scanning laser and photodetector array is 20 mm. Figure 3 (b)shows an experimental results of output-voltage of a PSD as a function of injected current to the laser A. Monotonic increase in the output is observed with the increase of injected current. The attached photographs are of the far-field pattern on the PSDs for each experimental condition. In this experiment, the separation of the lasers is 1200 μm.

4.Discussion

Inference speed of the beam scanning fuzzy controllers is supposed to be faster than 1M FLIPS (Fuzzy Logical Inference Per Second), because the scanning speed of the far-field beam of a Beam scanning laser diode has been measured to be less than 5 ns[5]. The bottle-neck of the speed is defuzzification process by a position sensing device. Currently the speed of the device made of Si is around sub-μs. The speed will be improved by miniaturization of the device and using higher-speed materials,such as GaAs.

The beam scanning fuzzy inference system is composed of fewer operational elements. The number of beam scanning laser diodes is the same as the sum of input-variables and result-fuzzy-functions. The number of photodetectors and PSDs are the same as the antecedent-rules and output-variables, respectively. This simple configuration results from parallel grade-evaluation and analog defuzzification by PSDs using the optical spatial parallelism. Because our reasoning method is different from conventional fuzzy reasoning methods, its feasibility was studied by numerically comparing the present and conventional methods. Figure 4 (a) and (b) shows a result of stability simulations of an inverted pendulum on a cart by Triangle-MIN-MAX-gravity method and by Gaussian-product-sum-gravity method, respectively. The figures show the number of inference reiterated by the time the pendulum stands stable, as a function of initial leaned angle of the pendulum. Parameter in the figures is the feedback gain of each controller. Smoother variation of and wider controllable range are realized in the case of Gaussian-product-sum-gravity method compared with conventional control methods.

5.Conclusion

Novel Fuzzy inference system using beam scanning laser diodes is proposed and experimental results are reported. This control system is simpler and its operation is estimated to be faster and smoother than conventional systems. These features come from the spatial parallelism of steered laser beams by beam scanning laser diodes.

6.References

1. E.H.Mamdani, "Applications of Fuzzy Algorithms for Control of Simple Dynamic Plant", Proc. IEE 121-12, pp.1585-1588 (1974).

2. G.C.Marsden, B.Olson, S.Esener, and S.H.Lee, "Optoelectronic Fuzzy Logic System", Technical digest of OPTICAL COMPUTING, Salt Lake City, pp.212-215 (1991).

3. S.Mukai, M.Watanabe, H.Itoh, H.Yajima, Y.Hosoi, and S.Uekusa, "Beam scanning and switching characteristics of twin-striped lasers with a reduced stripe spacing", Opt. and Quantum Electron., 17,pp.431-434 (1985).

4. M.Mizumoto, "Improvement of Fuzzy Controls (IV) -Case by Product-Sum-Gravity Method" (in Japanese language), Proc. 6th Fuzzy System Symp., Tokyo, pp.9-13 (1990).

5. S.Mukai, H.Yajima, S.Uekusa, and A.Sone: "Transverse second-order mode oscillations in a twin-stripe laser with asymmetric injection currents", Appl. Phys. Lett, 43, pp.432-434 (1983).

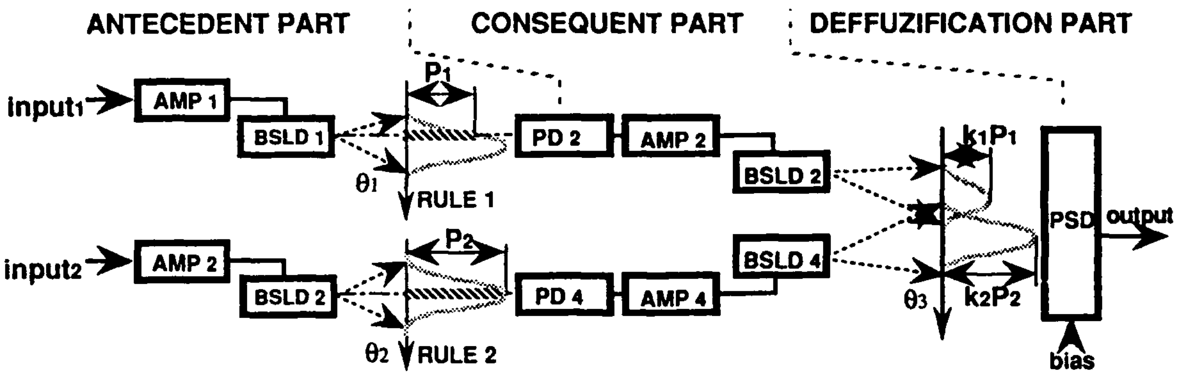
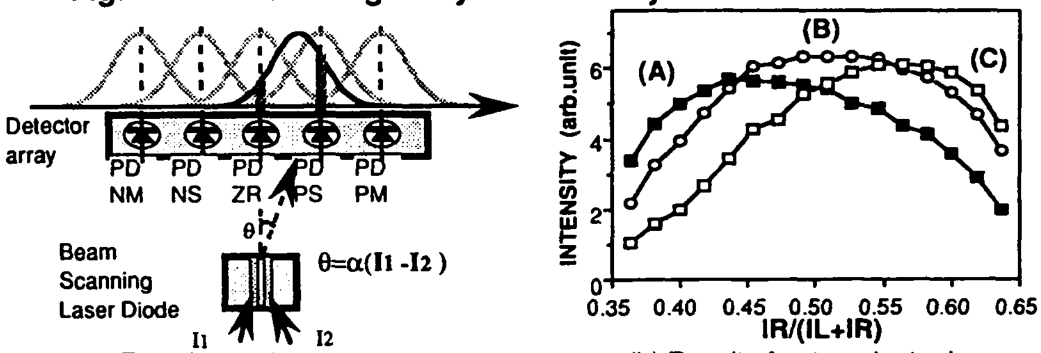
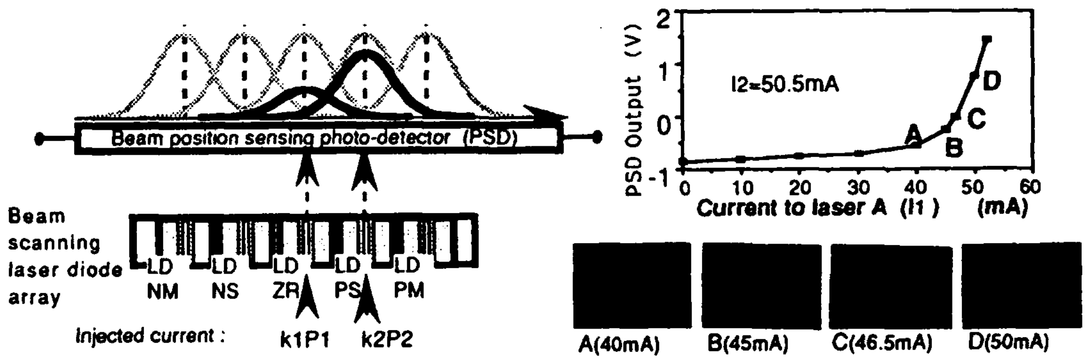


Fig.1 Beam scanning fuzzy inference system



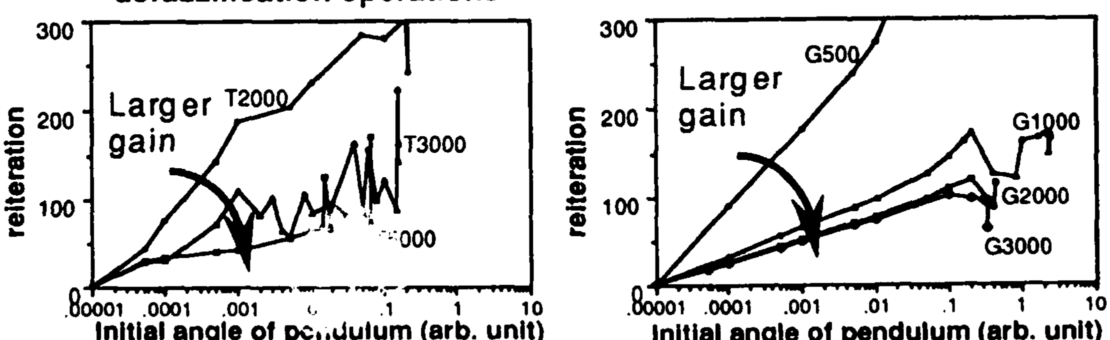
(a) Experimental setup (b) Result of antecedent rules

Fig.2 Experimental results of grade evaluation of antecedent rules



(a) Experimental setup (b) Defuzzifised result

Fig.3 Experimental results of consequeint, unification and defuzzification operations



(a) triangle-MIN-MAX-Gravity (b) Gaussian-product-sum-Gravity

Fig.4 Stability simulation of an inverted pendulum

Fourier Transforms of Fractional Order and their Optical Interpretation

David Mendlovic
Tel-Aviv University, Faculty of Engineering, 69978 Tel-Aviv, ISRAEL
Tel:972-3-6408245, Fax:972-3-6429540

Haldun M. Ozaktas
Electrical Engineering, Bilkent University, 06533 Bilkent, Ankara, TURKEY

Adolf W. Lohmann
Angewandte Optik, Staudtstr. 7/B2, 8520 Erlangen, GERMANY

The ath power of x, denoted by x^a, might be defined as "the number obtained by multiplying unity a times with x". Thus $(2.5)^3 = 2.5 \times 2.5 \times 2.5$. This definition makes sense only when a is an integer. However, it is an elementary fact that the definition of the power of a number can be meaningfully and consistently extended to real and even complex values of a. Likewise, the original definition of the derivative of a function makes sense only for integral orders, i.e. we can speak of the first or second derivative and so on. However, it is possible to extend the definition of the derivative to noninteger orders by using an elementary property of Fourier transforms. Bracewell shows how fractional derivatives can be used to characterize the discontinuities of certain *functions* [1]. An example from the field of optics is related to the Talbot effect [2], in which self-images of an input object are observed at the 2-D planes $z = Nz_0$ for integer N (z is the axial coordinate and z_0 a characteristic distance). Using a self-transformation technique [3], it was shown that N could also take on certain rational values.

In this letter we define fractional Fourier transformations in a similar spirit. The ath Fourier transform of a function $f(x, y)$ will be denoted as $\mathcal{F}^a[f(x, y)]$, or simply $\mathcal{F}^a f$ when there is no room for confusion. We require that our definition satisfy two basic postulates. First, $\mathcal{F}^1 f$ should be the usual first Fourier transform, defined as

$$(\mathcal{F}^1 f)(x', y') = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y)\, e^{-i2\pi(x'x+y'y)/s^2}\, dx dy, \tag{1}$$

where s, x, x', y, y' all have the dimensions of length. In a conventional "2f" optical Fourier transforming configuration [4], x, y would denote the coordinate of the input plane, x', y' those of the Fourier plane, and $s^2 = \lambda f$ (λ = wavelength of light, f = focal length of lens).

Our second postulate is to require that

$$\mathcal{F}^a[\mathcal{F}^b f] = \mathcal{F}^a \mathcal{F}^b f = \mathcal{F}^b \mathcal{F}^a f = \mathcal{F}^{a+b} f. \tag{2}$$

Consistent with our two postulates, $\mathcal{F}^{1/Q}[\cdot]$ may be defined for integer Q as that operation which when applied Q times gives the first (conventional) Fourier transform of f. An optical system which performs this operation may be realized by inserting a lens of appropriate focal length midway between the appropriately spaced input and output planes. It is now possible to define Fourier transforms of rational order $\mathcal{F}^{P/Q} f$ by repeated application of the operator $\mathcal{F}^{1/Q}$ for P times. The definition can be generalized to real orders by a limiting process.

Here we will not pursue this line of thought. Instead, we recall that the optical Fourier transforming operation is a result of the joint action of the dual operations of free-space propagation and focusing [5] [6] [7]. In a conventional "2f" system, the focusing action is concentrated at the lens location. The same operation can be performed by Q fractional Fourier transform stages in cascade each performing $\mathcal{F}^{1/Q}[\cdot]$. In this latter system, the act of focusing is evenly distributed through the act of propagation. In the limit that $Q \to \infty$, focusing and propagation will be infinitesimally and uniformly interspersed between each other. Of course, bulk systems with even moderately large Q would be quite impractical. Fortunately, systems satisfying this property can be realized as quadratic graded index (GRIN) media. Such media can be thought to consist of infinitesimal layers in which focusing and propagation take place simultaneously. The refractive index distribution in such a medium is given by [8]

$$n^2(r) = n_1^2(1 - (n_2/n_1)r^2), \tag{3}$$

where $r^2 = x^2 + y^2$ is the radial distance from the optical axis and n_1, n_2 are the GRIN medium parameters. By solving the ray equation, it was shown [8] that a parallel bundle of rays will be focused a distance $L \equiv (\pi/2)\sqrt{n_1/n_2}$ away from the input plane. If a function $f(x,y)$ is presented at the input plane $z = 0$, at the plane $z = L$ we observe $\mathcal{F}^1 f$ as given by Eq. 1 [9]. (This confirms that $z = L$ is the focal plane not only from the ray optics point of view but also from the physical optics point of view.) Now, since the system is fully uniform in the axial direction, $\mathcal{F}^a f$ can be physically defined as the functional form of the scalar light distribution at $z = aL$.

Above, we have motivated and defined the fractional Fourier transform in physical terms. However, it is important to note that fractional Fourier transforms can be defined purely mathematically and GRIN media introduced afterwards as a physical interpretation. We now present such a mathematical definition, also showing its relation to the physical definition above.

The self-modes of quadratic GRIN media are the 2-D Hermite-Gaussian (HG) functions [8], which form an orthogonal and complete basis set. The (l,m)th member of this set is expressed as

$$\Psi_{lm}(x,y) = H_l\left(\frac{\sqrt{2}x}{\omega}\right) H_m\left(\frac{\sqrt{2}y}{\omega}\right) \exp\left(-\frac{x^2+y^2}{\omega^2}\right), \tag{4}$$

where H_l and H_m are Hermite polynomials of orders l and m respectively, $\omega = (2/k)^{1/2}(n_1/n_2)^{1/4}$ with $k = 2\pi n_1/\lambda$, and λ is the wavelength. Each HG mode propagates through the GRIN medium with a different propagation constant [8]

$$\beta_{lm} = k\sqrt{1 - \frac{2}{k}\sqrt{\frac{n_2}{n_1}}(l+m+1)} \approx k - \sqrt{\frac{n_2}{n_1}}(l+m+1). \tag{5}$$

Any 2-D function $f(x,y)$ can be expressed in terms of the HG basis set as

$$f(x,y) = \sum_l \sum_m A_{lm}\Psi_{lm}(x,y), \tag{6}$$

$$A_{lm} = \int_{-\infty}^{\infty} f(x,y)\Psi_{lm}(x,y)/h_{lm}\, dx dy, \tag{7}$$

where $h_{lm} = 2^{-(l+m)}l!m!/\pi$.

Now, the fractional Fourier transform of $f(x,y)$ of order a can be defined as

$$\mathcal{F}^a[f(x,y)] = \sum_l \sum_m A_{lm}\Psi_{lm}(x,y)\exp(i\beta_{lm}aL). \tag{8}$$

It was shown [9] that our two postulates are satisfied by this definition, and that the scaling factor s appearing in Eq. 1 is given by $s = \omega/\sqrt{2}$. In the same references, we also discuss and prove some of the properties of fractional Fourier transforms, such as linearity, self-imaging parameters, intensity shift-invariance (as with the common Fourier transform), and generalize to complex values of the order a (which can be physically realized by attenuating media). We also define fractional convolutions (or correlations) through the equation

$$\mathcal{CONV}^a[f,g] = \mathcal{F}^{-a}[\mathcal{F}^a f \times \mathcal{F}^a g]. \tag{9}$$

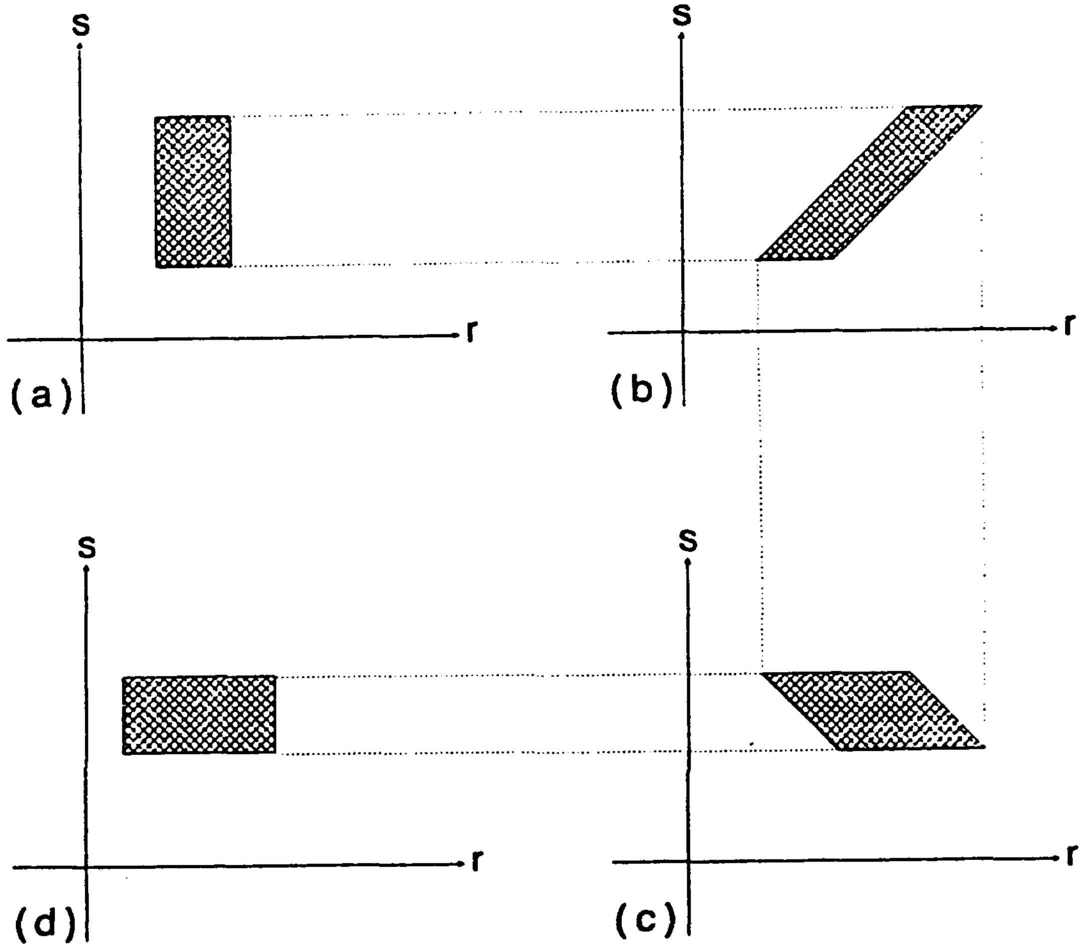


Figure 1: (a) Phase space representation of the original bundle of rays. (b) After free space propagation through a distance f. (c) After passage through a lens of focal length f. (d) After another free space propagation through a distance f.

To gain some additional insight regarding fractional Fourier transforms and also show its relation to ray optics, we now provide a phase-space interpretation. Let a particular paraxial ray be characterized by its radial distance r and slope s, both with respect to the optical axis at a particular axial position z. Then, the effect of passing through any optical system on this ray can be described by a movement in r-s space. For instance, free-space propagation corresponds to a horizontal displacement whereas focusing by a lens corresponds to a vertical displacement. Let us now consider a bundle of rays with a uniform spread of r and s (represented by the shaded rectangular region in Fig. 1a) and consider how this ray bundle is transformed as it passes through a conventional "2f" Fourier transforming configuration (Fig. 1). The overall effect of the "2f" system is to rotate the rectangular region by 90°, although the intermediate steps result in shearing of the rectangular region.

How do things look like in phase space if we use quadratic GRIN media instead? It is known that r and s obey the following equations in such media [8]:

$$\begin{aligned} r(z+\Delta z) &= r(z)\cos(\pi\Delta z/2L) - s(z)\sin(\pi\Delta z/2L) \\ s(z+\Delta z) &= r(z)\sin(\pi\Delta z/2L) + s(z)\cos(\pi\Delta z/2L), \end{aligned} \tag{10}$$

from which we can conclude that the region representing any given bundle of rays in phase space is uniformly rotated as we go from $z = 0$ to $z = L$. This uniform behavior is to be contrasted with that of the bulk "2f" Fourier transformer in which the focusing is concentrated at the lens, instead of being uniformly distributed throughout the system. This has discouraged us from basing the definition of fractional Fourier transforms on conventional bulk Fourier transformers. Although we cannot exclude the possibility of other definitions consistent with our two postulates, our particular definition is seen to be a natural and meaningful one, especially in an optical context.

Fractional Fourier transforms can form the basis of generalized spatial filtering operations, extending the range of operations possible with optical information processing systems. Conventional Fourier plane filtering systems [4] are based on a spatial filter introduced at the Fourier plane. This limits the operations achievable to linear space-invariant ones (i.e. operations which can be expressed as a convolution of the input function with a space-invariant impulse response). By introducing several filters at different fractional Fourier planes, it is possible to implement a wider class of operations. Note that full space-variant operations can be implemented using approaches such as (or equivalent to) vector-matrix multiplier architectures [10] or multi-facet architectures [11], but these result in a heavy penalty in terms of space-bandwidth product utilization. A detailed treatment of how many filters are needed and how they should be synthesized to realize a given operation must be postponed to another publication. Here we must satisfy ourselves by noting that this process, which in its extreme constitutes a generalization of conventional planar spatial filtering to volume spatial filtering, would ultimately be limited by noise and scatter from the several spatial filters.

It is important to note that although they provide a context for defining fractional Fourier transforms in a satisfying manner, GRIN media may not be a good candidate for their practical implementation, especially because they cannot have large space-bandwidth products. For practical purposes, it is possible to simulate GRIN media by using bulk lenses such that the same light distribution will be observed at the planes in which filtering will take place [9].

References

[1] R. Bracewell, *The Fourier Transform and Its Applications*, (McGraw-Hill, New York, second edition, 1986).

[2] K. Patorski, "The self-imaging phenomenon and its applications," in *Progress in Optics*, E. Wolf, ed. (North-Holland, Amsterdam, 1989), vol. 28 p. 3-110.

[3] A. Lohmann and D. Mendlovic, "Self-Fourier objects and other self-transform objects," JOSA **A9**, xxx-xxx (Oct. 1992).

[4] J. Goodman, *Introduction to Fourier Optics*, (McGraw- Hill, New York, 1968).

[5] A. Lohmann, "Ein neues Dualitatsprinzip in der Optik," Optik (Stuttgart) **11**, 478-488 (1954). In German.

[6] A. Lohmann, "Duality in Optics," Optik (Stuttgart) **89**, 93-97 (1992). A translation of Ref. 5.

[7] A. Vander Lugt, *Optical Signal Processing*, (Wiley, New York, 1991).

[8] A. Yariv, *Optical electronics*, (Holt Sounders Int Ed., third edition, 1985).

[9] D. Mendlovic and H. M. Ozaktas, "Fractional Fourier transformations and their optical implementation: Part I," JOSA A (submitted).

[10] J. W. Goodman, A. R. Dias and L. M. Woody, "Fully parallel, high- speed incoherent optical method for performing discrete Fourier transforms," Opt. Lett. **2**, 1-3 (1978).

[11] S. K. Case, P. R. Haugen and O. J. Løberge, "Multifacet holographic optical elements for wave front transformations," Appl. Opt. **20**, 2670-2675 (1981).

Optical Computing Posters: 2

OWE 8:00pm-9:30pm
Grand Ballroom Center

Optical transpose interconnection system

Philippe J. Marchand, Gary C. Marsden, and Sadik C. Esener

University of California, San Diego
Department of Electrical and Computer Engineering
La Jolla, Ca 92093-0407
Ph: 619-534-4943 Fax: 619-534-1225

1. Introduction

Optical interconnections have been shown to have advantages over electrical interconnections in terms of speed, energy and density for global links.[1,2,3] In addition, the flexibility of optical interconnections allows efficient electronic layouts which can improve the performance of electrical connections in an optoelectronic system. In this paper we are concerned with a particular one-to-one interconnection, the transpose interconnection. The Optical Transpose Interconnection System (OTIS) is a simple, efficient, and scalable means of providing a transpose operation utilizing only a pair of lenslet arrays. The usefulness of the transpose interconnection can be shown for several optoelectronic architecture classes [4,5]. This paper briefly reviews the transpose interconnection functionality and describes some of the architectures supported by OTIS. The OTIS optical system and various scalability parameters are analyzed including system length and volume, system power efficiency, and crosstalk. Variations of the basic OTIS optical system for different applications are introduced: a generalized system that can match the geometry of any arbitrary optoelectronic chip layout, the folded, the bi-directional, and the multi-channel systems. Experimental results of a 64x64 channels OTIS and the design optimization of this experimental system are provided.

2. Transpose Interconnection

A transpose operation usually consists of symmetrically interchanging the elements of a matrix with respect to the first diagonal of that matrix. We propose to implement a particular transpose interconnection between two arbitrary planes where each row of the input and output arrays is arranged in a raster format within that array as shown on figure 1. The transpose interconnection we describe is a one-to-one interconnection between L transmitters and L receivers where L is the product of two integers, M and N. Such an interconnection is called an MxN transpose.

Figure 1 shows the physical layout of the transmitter and receiver planes of OTIS for M=16 and N=4. Both planes are divided into regions in which one of the indices is constant, the other index showing the location of the node within that region. The gray transmitters and receivers on figure 1 show a pattern of signals mapped from the transmitter plane to the receiver plane via the transpose interconnection.

Figure 2 shows the side view of the OTIS optical system that connects the two planes described above; the top view would be similar. A $\sqrt{N}x\sqrt{N}$ array of lenslets is placed in front of the transmitter plane and a $\sqrt{M}x\sqrt{M}$ array of lenslets is located before the receiver plane. The interconnection from transmitters to receivers is traced by the chief ray passing through the center of the transmitter lens and the center of the receiver lens. These two lenses comprise an imaging system between the transmitter and receiver planes. Note that the optical imaging system requires that the indexing in the receiver plane be rotated 180 degrees relative to the transmitter plane.

3. Optical Transpose Interconnection System Architectures

An important application of OTIS is in support of multistage interconnection network architectures (OTIS-MIN) based on k-shuffles.[6] A k-shuffle MIN functionally provides full connectivity between L input channels and L output channels in $\log_k L$ stages of optoelectronic switch planes and $(\log_k L)-1$ stages of optical k-shuffles. One optoelectronic switch has k optical inputs (receivers) and k optical outputs (transmitters) and provides crossbar equivalent electronic switching between its k channels. The k-shuffle is equivalent to a (L/k)xk transpose. An interesting application of OTIS-MIN arises when $k=(MN)^{1/2}$ since the number of stages for routing becomes a constant: $\log_k L = \log_{\sqrt{MN}} MN = 2$. In this case, only 2 stages of optoelectronic switches and 1 stage of optics are required to perform full routing between source and destination planes, independent of the total number of communication channels in the MIN.

Another useful application of OTIS is in support of the two-dimensional mesh-of-trees architecture (OTIS-MT), which is useful for matrix computations.[5] The OTIS-MT layout is similar to that required for the D-STOP architecture, which was also designed for matrix algebraic computations.[7] The primary physical difference is that the transmitters in the OTIS-MT architecture reside in the leaf units of the tree, whereas the transmitters of D-STOP reside in the root units. In the D-STOP vertical distribution,[7] the broadcast of N signals to MN receivers is achieved through N transmitters, each with an optical fanout of M. The transpose interconnection of the OTIS-MT architecture is functionally equivalent to the D-STOP vertical distribution, except that the broadcast is achieved by first distributing each signal through an electronic fanout to M transmitters. While there are M times as many transmitters in the OTIS-MT architecture as in D-STOP, the required optical fanout is 1 due to the one-to-one nature of the transpose interconnection. Whereas the large optical fanout requirement of D-STOP limits the choice of transmitter technologies it might utilize, no such limit is imposed on the OTIS-MT architecture.

Often, interconnection patterns have recognizable subclasses of similar, relatively simple interconnections, but the necessary locality of transmitters or receivers from different subclasses makes it difficult to exploit this piece wise simplicity. OTIS can be used in some cases to spatially isolate these subclasses into regions where additional simple optical or electronic transforms can be implemented. As an example, it can be shown that OTIS can be used as a component of a hypercube interconnection.

4. Optical Transpose Interconnection Systems

The layouts and the optical system described previously assume that transmitters and receivers all lay on a regular square grid. In practice, due to the electronic layout constraints, it could prove useful for the optical interconnection system to be able to accommodate optoelectronic chips with different arrangements. Figure 3 shows one possible case where there is a gap between some of the nodes in both the transmitter (C_t) and receiver (C_r) planes and where the spacing between nodes in the transmitter plane (Δ_t) is different from the spacing between nodes in the receiver plane (Δ_r). The OTIS optical system still performs the transpose interconnection without modifications. Note that these on-chip gaps can be used for wiring electrical inputs and outputs on and off the chips, for bringing power lines to the different locations on the chips, or to accommodate the fact that the optoelectronic transmitter and receiver planes are built using multi-chip module technology.

The systems in figure 2 and 3 perform the transpose interconnection but leave the transposed result on a different array than the original one. In some cases, it may be desirable to have both the original matrix and its transpose on the same array. OTIS can be easily adapted to accommodate these requirements. A single imaging lens and a folding mirror in a one-to-one, 4f configuration are simply added to a single plane of OTIS lenses. The interconnection function achieved in this case is exactly that of a transpose operation without any image inversion or rotation, leaving the elements of the first diagonal of the matrix unchanged and interchanging all the other elements of that matrix with respect to the first diagonal. It is important to note that only symmetrical systems (N=M) can be folded in this manner.

An important feature of the Optical Transpose Interconnection System described in the previous sections is that it can be made bi-directional very easily. This could prove useful in a number of applications. For interconnection networks in a packet switching configuration it allows blocking information to be sent back to the previous stages of the network in order to establish the correct routing path in the network. OTIS can also be used in a matrix-vector multiplier configuration (OTIS-MT) in support of neural network applications. In this case, the bi-directionality characteristic of OTIS can be used to implement an on-chip fully parallel back propagation learning system.

A final interesting feature of the OTIS system is that it can become a multi-channel system where each node has more that a single pair of transmitter/receiver. This is useful in the case where more than bit serial communication per channel is required in a system. It may also prove useful in the case of communication systems where one transmitter/receiver pair is used for data transfer and one other for activation/status information. Finally, it might be required to have 2 transmitter/receiver pair for each communication channel in a system in order to implement differential readout (dual-rail logic) in the case where the intrinsic signal-to-noise ratio of the system is not sufficient to provide low communication bit error rates.

5. Experimental Results

The experimental system demonstrates the viability and feasibility of the OTIS concept by implementing a one stage 64x64 (NxM) symmetrical (N=M) OTIS optical system. The system is made of a 64x64 pinhole array, 2 refractive lenslet arrays, and a CCD camera as the output plane detector. The 64x64 pinhole array is fabricated in-house using our Electron beam lithography system. The pinholes have a 10x10 μm square aperture and they are spaced 57 μm apart for a total input plane size of 3.64 mm. As illustrated in figure 4, the pattern formed by these pinholes (input array) represents the letters OT. The lenslet arrays were purchased from Adaptive Optics Associates and are fabricated in stamped epoxy on a glass microscope slide. The lenslet array size is 20x32 and each lenslet has a 399 μm aperture for a focal length of 3.2 mm leading to a $f_\#$ of approximately 8. The total length of the system is measured to be around 37 mm. As expected, the system achieves a perfect one to one imaging and the output plane pattern (Figure 4) consists of the OT letters reduced in size by a factor of 8 ($\sqrt{N}$) and replicated 8x8 times over the output plane. Note that the actual output plane (figure 4) shows more than 8x8 replications of the input pattern since the lenslet arrays contain more than 8x8 lenslets.

6. Conclusions

OTIS is useful for various classes of optoelectronic architectures and in particular for shuffle-based interconnections and as a component of a hypercube system. The basic OTIS optical system analysis is based on various scalability parameters including system length, volume, and efficiency and shows that scalability is not limited by diffraction. Variations of the basic OTIS optical system for different applications can be shown to be useful, such as the generalized system that can match the geometry of any optoelectronic chip layout, as well as the folded, bi-directional, and multi-channel systems. The experimental demonstration of a 64x64 channels OTIS has been successfully performed. Finally, aberrations in an OTIS optical system based on diffractive optical elements can be virtually eliminated using Code V[8] to design aspheric holographic optical elements.

References

1 M. Feldman, S. Esener, C. Guest, and S. H. Lee, *Applied Optics* **27**, 1742-1751 (1988).

2 M. Feldman, C. Guest, T. Drabik, and S. Esener, *Applied Optics* **28**, 3820-3829 (1989).

3 F. Kiamilev, P. Marchand, A. Krishnamoorthy, S. Esener, and S. H. Lee, *Journal of Lightwave Technology* **9**, 1674-1692 (1991).

4 G. Marsden, P. Marchand, and S. Esener, *OSA Annual Meeting*, Albuquerque, September 1992.

5 F. T. Leighton, Introduction to Parallel Algorithms and Architectures, Morgan Kaufmann (1992).

6 A. Krishnamoorthy, P. Marchand, F. Kiamilev, and S. Esener, *Applied Optics* **31**, 5480-5507 (1992).

7 G. Marsden, A. Krishnamoorthy, S. Esener and S. H. Lee, *Optics Letters* **16**, 1970-1972 (1991).

8 Code V is registered trademark of Optical Research Associates, Pasadena, Ca.

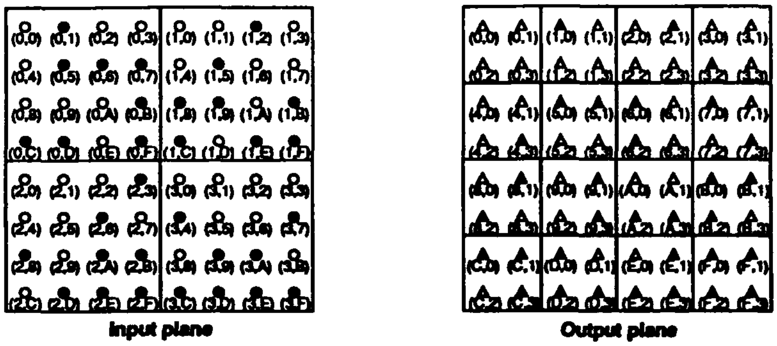


Figure 1: Transmitter and receiver planes showing the mapping of the OTIS interconnection (shaded sources and detectors)

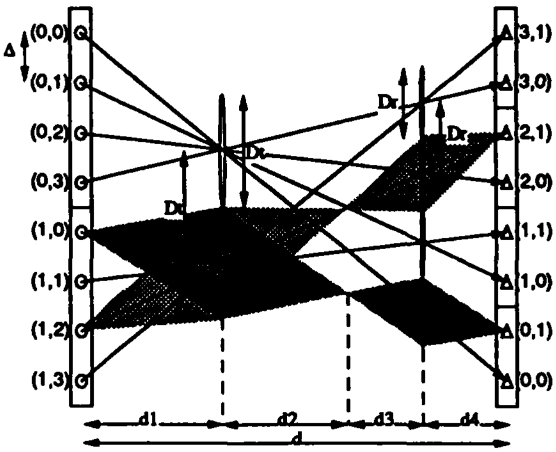


Figure 2: Side view of the OTIS optical system showing one-to-one imaging through coupled lenslet arrays

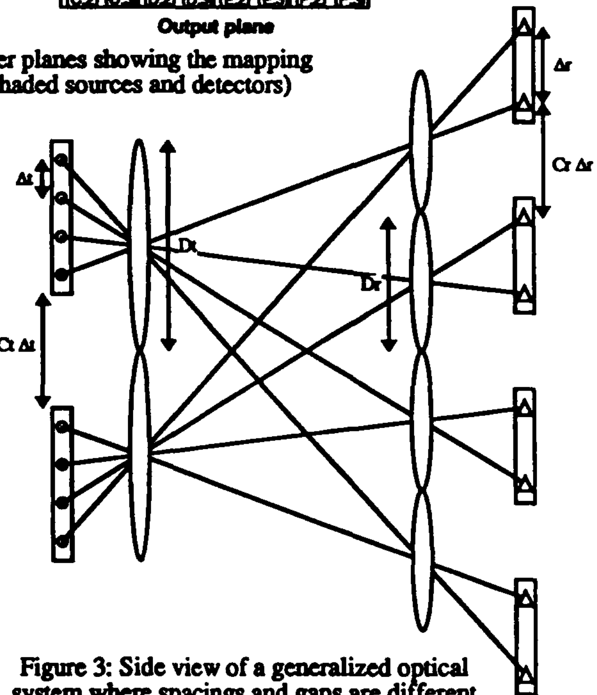
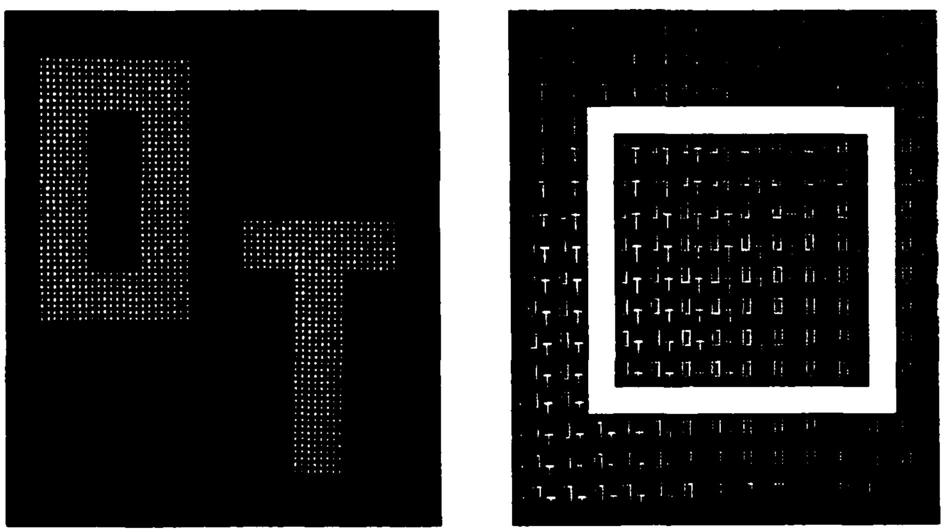


Figure 3: Side view of a generalized optical system where spacings and gaps are different between transmitter and receiver planes



Input plane

Output plane

Figure 4: Experimental demonstration of a 64x64 OTIS optical system. The input plane consists of a 64x64 array of pinholes forming the letters OT and the output plane consists of these same letters reduced by a factor of 8 and replicated 8x8 times. The output plane actually shows more than 8x8 patterns since oversized (20x32) lenslet arrays were used for the experiment.

Theoretical study of the operating conditions for Symmetric Self-Electrooptic-Effect-Devices within cascaded digital optical arrays

Marc P.Y. Desmulliez, John F. Snowdon
and Brian S. Wherrett

Department of Physics, Heriot-Watt University,
Riccarton, Edinburgh, EH14 4AS, Scotland, UK
Tel : UK 31 451 3048 Fax : 3136

Considerable interest has been focused over the last few years on use of the symmetrical self-electro-optic-effect-device or S-SEED[1-3]. Two dimensional arrays of S-SEEDs have been exploited in demonstrator digital optical circuitry[4] and photonic-switching fabrics[4-6]. An S-SEED version of the optical-cellular-image processor module has also recently been realized[7]. These circuits and other reported measurements[8] demonstrate the cascability of the devices and their robustness to fluctuations in optical and device parameters. Having in mind the use of S-SEED arrays in more complex architectures we have been concerned to quantify the degree of robustness - the tolerance - of S-SEED components.

In this paper we will present the results of a tolerancing analysis based on a methodology applied earlier to nonlinear-interference-filter optical logic components[9]. We will account for the leakage of optical power from previous S-SEEDs in an iterative architecture, onto the device under analysis ; this leakage is inherent to the manner in which S-SEED arrays have been operated to date. Non-uniformity of beams within an array, or of the devices themselves, is also treated.

In general the tolerance conditions produce volumes of acceptable device operation within the multi-dimensional space of the parameters describing the device. In particular we concentrate on the parameters that are adjustable, once a given set of devices are interfaced together within a circuit. The target is therefore to provide a recipe for the successful construction and operation of logic circuits. Such a recipe is essential for complex circuits which exhibit many adjustable parameters.

We develop an empirical model for the current-voltage and reflectivity voltage characteristics of the SEED device[1-3] which gives the steady states and the dynamic behaviour of the symmetric-SEED[10]. The model has been chosen to be sufficiently simple so that tolerance analysis can be carried out without excessive computational demands, but it is also sufficiently realistic so that close fits between the theoretical optical responses and those measured experimentally can be made. The circuit considered consists of cascaded two-dimensional arrays of S-SEEDs used as programmable optical logic gates such as in the existing optical cellular-logic-image-processor, the O-CLIP[7].

The device performance is analysed in the cases of perfect balance (2-D uniformity) and inbalance of the various beams. Worst-case conditions of operation determine where the device will work correctly in the multi-dimensional space generated by the relevant adjustable parameters[9]. The volumes of acceptability then created enable us to calculate the tolerance limits. In the perfect balance case, the analysis includes the leakage effects from the previous arrays.

Three different clocking procedures are taken into account, differing by the manner in which preset and control beams are used to determine the logic functionality of the gates.The relative merits of the procedures will be described. Within the regions of acceptability we are also able to determine the latency-time for circuit operation. The trade-off between bandwith and tolerance levels can therefore be discussed.

In figures 1 and 2 are shown the volumes of acceptability of the S-SEED under one clocking procedure. In the balanced operation (figure 1), the parameter space is described by three independent combinations of the adjustable parameters: the applied voltage V_0, the inter-gate gain G and the preset over the read-beam power-ratio P_p/P_r. The axis generated by the last parameter is on a logarithmic scale to emphasize the large fluctuations tolerable in that ratio. The labels on the V_0 axis refer to specific values employed by the empirical model and in this example, gain values less than 0.5 are relevant since a 1-D nearest neighbour optical interconnect is considered. The volume shows large tolerance in all parameters. In the case of unbalanced operation (figure 2) applied voltage V_0, window- and route-dependent power ratios are the chosen adjustable parameters. The constant-voltage cross-sections of the volume can be approximated by rectangles of vertex coordinates (1,[cT-1]/[c-T]), (1,[c-T]/[cT-1]), (T,1), (1/T,1) where c and T are respectively the output contrast ratio and the threshold. These sections exhibit large tolerances at high voltages.

The paper will detail the S-SEED modelling, the methodology and the results of the tolerance analysis and indicate the extension of the method to alternative circuits such as those incorporating logic-SEED arrays. The method differs from earlier ones[11] in that leakage effects are considered and dependence on applied voltage, as well as the non-uniformity. Leakage effect is inherent to the manner in which circuits have been operated and can dominate the tolerance combinations.

References

1. D.A.B. Miller and al, "The quantum well self-electro optic effect device : Opto-electronic bistability and Oscillation and Self-Linearized Modulation", IEEE J. Quantum Electron. **21** , 1462-1475 (1985).

2. D.A.B. Miller, "Quantum well self electro-optic effect devices", Opt. and Quant. Elect., **22** , pp 561-598, (1990).

3. A.L. Lentine, H.S. Hilton, D.A.B. Miller, J.E. Henry, J.E. Cunningham and L.M.F. Chirovsky, "Symmetric self electro-optic effect device : optical set-reset latch, differential logic gate and differential modulator/detector ", IEEE J. Quant. Elect., QE **25** , pp 1928-1936, (1989).

4. M.E. Prise, N.C. Craft, R.E. LaMarche, M.M. Downs, L.A. D'Asaro, L.M.F. Chirovsky, "Cascaded operation of arrays of symmetric self electro-optic effect devices", Appl. Opt., **30** , pp 2841-2843, (1991).

5. F.A.P. Tooley and F.B. McCormick, "Design issues in free space optical interconnection of switching systems", O.S.A. Proceedings on Photonic Switching, **8** , pp 38-42, (1991).

6. F.B. McCormick, F.A.P. Tooley, T.J. Cloonan, J.L. Brubacker, A.L. Lentine, R.L. Morrison, S.J. Hinterlong, M.J. Herron, S.L. Walker, J.M. Sasian, " S-SEED based photonic switching network demonstration", ibid., pp48-55, (1991).

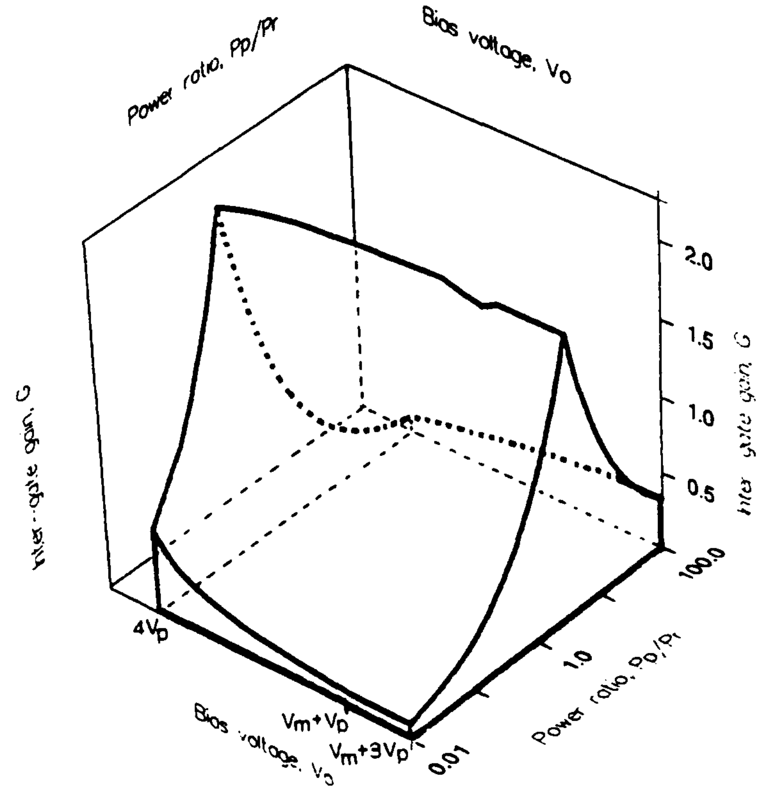
7. F.A.P. Tooley, S. Wakelin, "Design of a Symmetric Self Electro-Optic Effect Device Cellular Logic Image Processor", accepted for publication in Appl. Opt. (1992).

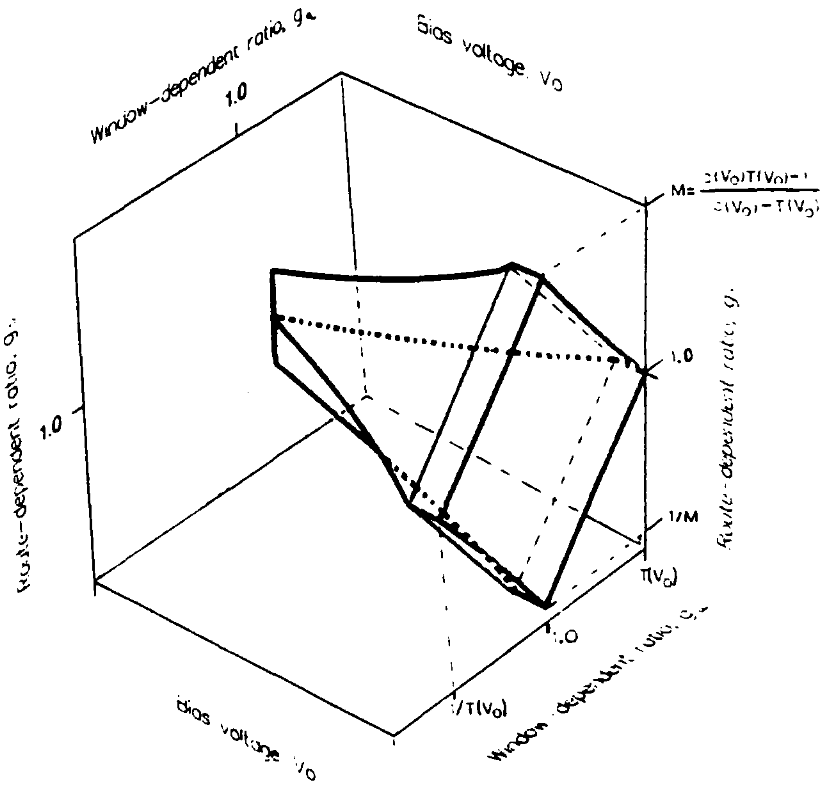
8. L.M.F. Chirovsky, M.W. Focht, J.M. Freund, G.D. Guth, R.E. Leibenguth, G.J. Przybylek, L.E. Smith, L.A. D'Asaro, A.L. Lentine, R.A. Novotny, D.B. Buchholz, "Large arrays of SEEDs", ibid., pp 56-59, (1991).

9. B.S. Wherrett and J.F. Snowdon, "Tolerance studies for digital optical computing circuitry with application to nonlinear interference filters", International J. Opt. Computing, **1** , pp 41-70, (1990).

10. B.S. Wherrett, M.P.Y. Desmulliez, J.F. Snowdon, "Operating conditions for Self-Electrooptic-Effect-Devices within digital optical circuits", to be submitted in Optical Computing and processing.

11. S. Wakelin and F.A.P. Tooley,"The tolerance to spatial noise of differential logic gates", Optical Computing and processing, vol. 2, **2** , (1992).





Figures 1,2 : Volumes of acceptability of an S-SEED in balanced and unbalanced operations.

Reversal input superposing technique for all-optical neural networks

Yoshio Hayasaki, Ichiro Tohyama, Toyohiko Yatagai
Institute of Applied Physics, University of Tsukuba, Tsukuba, Ibaraki 305
Tel:+81-298-53-5317

Masahiko Mori and Satoshi Ishihara
Electrotechnical Laboratory, Tsukuba, Ibaraki 305
Tel:+81-298-58-5623

1. Introduction

Basic operations of a simple neural network model are a spatial weighted sum operation, including arithmetic operations of multiplication and addition, and a nonlinear operation. A neuron calculates multiplications of synaptic weights and input signals from other neurons, and performs the nonlinear operation for their sum. In a general neural network model, because the synaptic weights have the bipolar values, which are excitatory weights and inhibitory weights, the arithmetic operations need addition, subtraction, and multiplication in a negative range. To represent the bipolar values, a spatial position encoding technique[1,2] and a polarization encoding technique[3,4] are applied. These techniques representing the bipolar values require twice elements of original neural networks and subtraction of signals from a pair of elements encoded to negative and positive values. The optical neural networks using these techniques perform subtraction with electric circuits, and an optoelectronic signal conversion is required. An optical neural network model using the neurons encoded to positive and negative was proposed.[5] This model requires twice number of the neuron.

In this paper, a new technique for the optical neural networks is presented. This technique can perform the neural computation in positive range and the number of elements is same as original neural networks. The bipolar weights are changed to the non-negative weights by adding a constant. By superposing of reversal inputs with weighted sums, subtraction in the neuron is a constant. This means that subtraction is not necessary, because subtraction with the constant can be treated as an offset of a nonlinear output function. This technique is called the reversal input superposing technique (RIST) which is obtained from modifying a basic calculation rule of the neural networks, and implemented by a conventional configuration of optics. We construct an optoelectronic system for the experimental verification of the present technique. Furthermore, we demonstrate the Hopfield model[6] as an example of the neural network model realized by the experimental system.

2. Modification of Basic Neural Network Model

Consider a discrete model of simple neural networks consisting of M neurons, which receive the same input signals $X = \{x_1, ..., x_i, ..., x_N\}$ from N input neurons, and emit respective output signals $V = \{v_1, ..., v_j, ..., v_M\}$. Let w_{ji} be the synaptic weight from the i-th input neuron to the j-th output neuron. The j-th weighted sum of the input signals is then written as

$$u_j = \Sigma_i w_{ji} x_i - h_j, \qquad (1)$$

where h_j is the j-th value of a set of the thresholding values $H = \{h_1, ..., h_j, ..., h_M\}$. The output v_j is written as

$$v_j = f(u_j), \qquad (2)$$

where f is a nonlinear output function. The synaptic weight w_{ji} and the weighted sum u_j are generally the finite real values. The input x_i and the output v_j are ranged from 0 to 1 in many neuron models.

Here, considering to introduce two constants α and β, Eq. (1) is rewritten as

$$u_j = \Sigma_i(w_{ji} + \alpha)x_i + \alpha\,\Sigma_i(\beta - x_i) - N\alpha\beta. \qquad (3)$$

The thresholding values $H = \{ h_j \}$ are neglected because they can be treated as same as the synaptic weights. The constant α is a bias of the weights w_{ji}, and is determined to

$$\alpha > -\min(w_{ji}), \qquad (4)$$

where $\min(w_{ji})$ denotes the minimum value of a set of the weights $W = \{ w_{ji} \}$. Hence, the biased synaptic weights $W^b = \{ w^b_{ji} \} = \{ w_{ji} + \alpha \}$ of the first term in Eq. (3) are necessary to be positive. The constant β is introduced so that the second term in Eq. (3) is positive, and is usually set to

$$\beta > \max(x_i), \qquad (5)$$

where $\max(x_i)$ denotes the maximum value of a set of input signals $X = \{ x_i \}$. All the modified input signals $\overline{X} = \{ \overline{x}_i \} = \{ \beta - x_i \}$ of the second term in Eq.(3) are positive. Third term $N\alpha\beta$, which is constant since the constants α and β are determined in advance, can be treated as an offset value of the nonlinear output function f. The weighted sums $U = \{ u_j \}$ of Eq. (1) are rewritten by using the biased weights W^b and the modified input signals $\overline{X}$, and the modified weighted sums $\overline{U} = \{ \overline{u}_j \}$ are written as

$$\overline{u}_j = \Sigma_i w^b_{ji} x_i + \alpha\Sigma_i\overline{x}_i. \qquad (6)$$

The modified weighted sums $\overline{U}$ are non-negative values given the constant $N\alpha\beta$ to the weighted sums $U = \{ u_j \}$. The first term in Eq. (6) is the vector-matrix product between the biased weights W^b and the input signals X. The second term in Eq. (6) is the sum of the modified input signals $\overline{X}$. The output signals $V = \{ v_j \}$ of Eq. (2) are also written as

$$v_j = f'(\overline{u}_j), \qquad (7)$$

where

$$f'(x) = f(x - N\alpha\beta). \qquad (8)$$

The nonlinear operation $f'(x)$ is generally monotonic increasing. In the model of the neural computations of Eqs.(6), (7), and (8), the important feature is that all the synaptic weights and all the neural signals are non-negative values.

3. Optical Implementation of RIST

Equations.(6), (7), and (8) equivalent to the original neural computations of Eqs.(1) and (2) fit to an all optical implementation, because of no subtractions and no negative values. The modified input signals $\overline{X}$ in Eq. (6) are the reversal signals of the input signals X if the constant β is chosen to be 1. Accordingly, addition in Eq. (6) can be calculated by superposing the optical weighted sum with the reversal input signals. Equation (7) is based on non-negative calculation, and $N\alpha\beta$ in Eq.(8) is treated as an offset of the thresholding device.

The RIST requires simultaneously an input and a reversal input. As an example of the implementation of the RIST, we think about a polarization encoding method. By modulating the input signals to a polarization pattern, the input pattern and the reversal pattern are obtained in parallel. The real value from 0 to 1 is represented by the angle of the polarization from 0° to 90° with respect to the vertical axis. The polarization pattern changes to the light intensity pattern with an analyzer set to a vertical direction for obtaining the input pattern. The analyzer, on the other hands, is set to the horizontal direction for obtaining the reversal pattern.

In order to verify the RIST, the experimental system is constructed as shown in Fig.1. The system is composed of a liquid crystal television (LCTV), a photographic film, a CCD camera, a computer, an image processor, and other optical components. A light source is a He-Ne laser, which is omitted in Fig. 1. The LCTV, which is removed the analyzer, accepts the signals through the image processor from the computer and modulates the input signal to the polarization pattern. The light modulated by the LCTV is divided to two orthogonal polarized light by using Analyzer1 and Analyzer2, and two divided lights are superposed by the beam splitter. The angles of Analyzer1 and Analyzer2 are set to 0° and 90° with respect to the vertical axis, respectively. Analyzer3 is used to adjust a ratio of the intensities of two divided lights. The photographic film memorizes the synaptic weights W^b. The optical part in the experimental system calculates multiplications and additional in the summation of the following equation rewritten from Eq. (6),

$$\bar{u}_j = \Sigma_i(w^b_{ji}x_i + \alpha\bar{x}_i). \qquad (9)$$

The computer and the image processor in the electric part of the experimental system handles a summation of Eq.(9) and the nonlinear output function f of Eq.(8). The computer sums up the signals from the CCD camera, and calculates the output signals $V=\{v_j\}$ of the neurons by performing the nonlinear operation f to the summation.

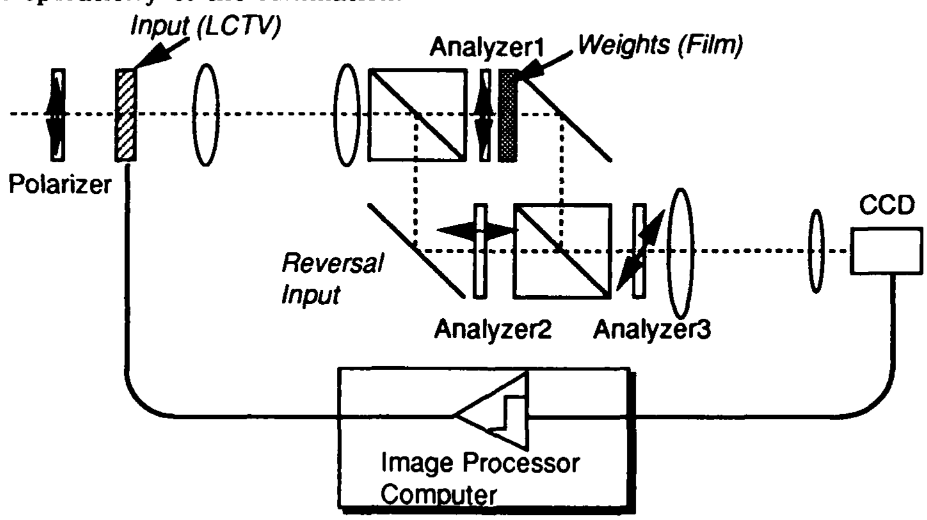


Figure 1 Experimental neural system based on RIST

4. Experiment

In the experimental system, the Hopfield model is implemented to made a fundamental verification of our technique. The synaptic weights $W = \{w_{ji}\}$ for storing a set of P binary vectors $X_p = \{x_{1p}, x_{2p}, \ldots, x_{Np}\}$ consisting N bits long are given by

$$w_{ji} = \Sigma_p(2x_{jp} - 1)(2x_{ip} - 1) \quad \text{for } i \neq j,$$
$$= 0 \quad \text{for } i = j. \qquad (10)$$

Because the minimum value of the synaptic weights W is -P, the biased weights $W^b=\{w_{ji}+P\}$ added the constant $\alpha=P$ are obtained. The offset of the nonlinear operation in Eq.(8) is set to be NP.

A Hopfield model with 25 neurons is implemented in the experimental system. Figure 2 shows three stored patterns consisting 5 x 5 elements. Figure 3 shows the results of the recalling for the retrieval of three stored patterns in the experimental system. The recallings in the experimental system were performed for all possible initial inputs whose deviation from the stored pattern are less than 3. For comparison, the results of the computer simulation of the recallings are shown. Figures 3(a), 3(b), and 3(c) show the recalling results for three stored patterns A, B, and C, respectively.

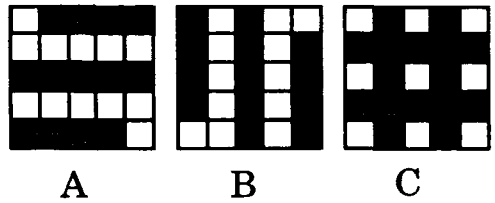


Figure 2 Three stored paterns

5. Conclusion

A general technique for optical neural networks called the reversal input superposing technique (RIST) has been discussed. Optical neural networks introducing the technique are not necessary negative weights and subtraction, and inherently constructed all optical system without the electric circuits. Also the required elements of the neuron and the weight is same number of the original neural networks. Therefore optical neural system is simple configuration by a conventional lens system and a masking of the image.

The RIST is requires the input pattern and the reversal pattern in parallel. They are obtained by modulating input signals to polarization pattern. In the experimental system, a LCTV is utilized. The experimental system performs the weighted interconnections and the superposition of lights on both optical paths in optical part, and the nonlinear output function in the electric part. In the experimental system, the Hopfield model which has 25 neurons and stores 3 patterns was implemented. In the experimental results, the convergency for the retrieval of three stored patterns is almost similar to the computer simulations.

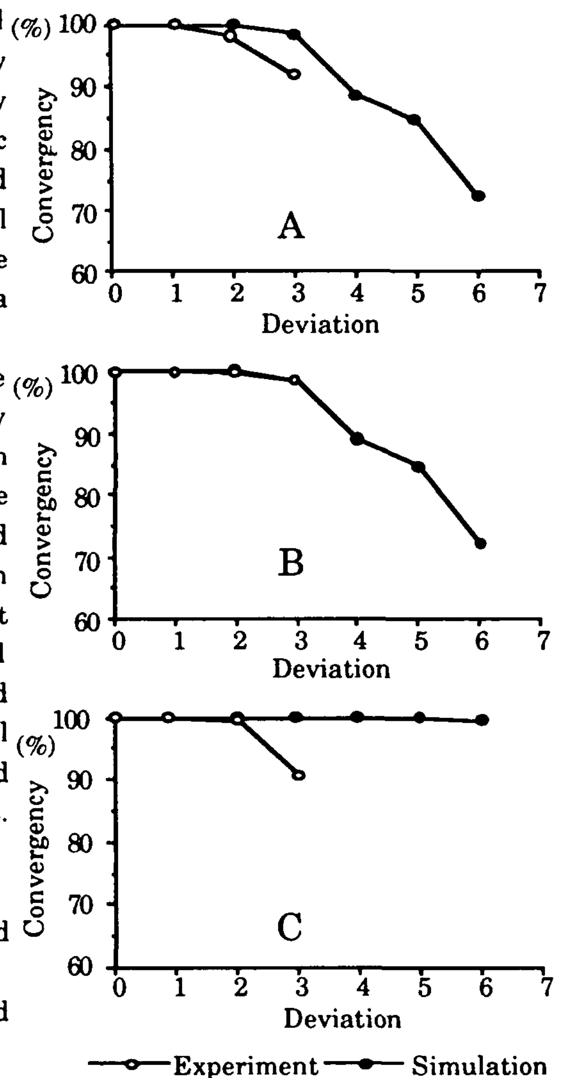


Figure 3 Expreimental results of recalling.

References

1. J. Ohta, S. Tai, M. Oita, K. Kuroda, K. Kyuma, and K. Hamanaka, Opt. Lett., **14**, 844 (1989).
2. N. Kasama, Y. Hayasaki, T. Yatagai, M. Mori and S. Ishihara, Jpn. J. Appl. Phys., **29**, L1565 (1990),.
3. M.Kranzdorf, B.J.Bibner, L.Zhang, and K.M.Johnson, Opt. Eng. **28**, 844-848 (1989).
4. I.Shariv, and A.A.Friesem, Opt. Lett. **14**, 485-487 (1989).
5. C.H.Wang and B.K.Jenkins, Appl. Opt. **29**, 2171-2186 (1990).
6. J. J. Hopfield, Proc. Natl. Acad. Sci. USA **79**, 2554 (1982).

Tandem D-STOP Architecture for Error Backpropagation Networks

Gary C. Marsden, Ashok V. Krishnamoorthy, Jean Merckle*, and Sadik C. Esener
Department of Electrical Computer Engineering
University of California, San Diego
La Jolla, CA 92093-0407
Tel: 619-534-1743
Fax: 619-534-1225
E-mail: gray@topcad.ucsd.edu
* Universite de Haute-Alsace, Mulhouse, France

Artificial neural networks have become popular, to a large extent, because of their putative ability to adaptively learn their optimal operational states. It is also generally accepted that parallel systems with on-line learning must be designed in order to satisfy the demand for real-time, large-scale networks. We have recently designed[1] and demonstrated[2] a scalable, prototype optoelectronic system which supports the recall, or feedforward, mechanism of neural networks. In this paper, we present a modification of this architecture which allows on-line, fully parallel learning.

The Dual-Scale Topology Optoelectronic Processor[3] (D-STOP) is a fully parallel optoelectronic architecture designed for matrix algebraic processing, which uses a minimum number of optical transmitters while maintaining area efficient electronics and simple, space-invariant optical interconnects. For an MxN matrix problem size, the architecture consists of M processing elements (PEs), one for each row of the matrix. Within each PE is an electronic binary tree structure with processing sub-units at each node. Each PE has N leaf units which correspond to the matrix elements within the given row. Each leaf unit has an optical detector, as well as local memory, logic circuitry, and electronic I/O to the tree structure. At the root of the tree is the root unit, which in addition to local memory, logic and electronic I/O has an optical transmitter for data output. Each PE receives a copy of the optical input vector, with each leaf receiving one of the N elements of this input vector.

Figure 1 shows how two D-STOP arrays can be coupled to implement a three layer neural network. The input vector elements $x_k(0)$ are optically fanned-out to each of the hidden (Array 1) layer neurons. The output $x_j(1)$ from each of these hidden layer neurons is given by,

$$x_j(1) = f\{\textstyle\sum_k W_{jk}(1) \times x_k(0)\} \qquad [1a]$$

The product $W_{jk}(1)\times x_k(0)$ is calculated at the k-th leaf unit of the j-th PE. The summation of these products within a PE is achieved through electrical fan-in on the binary tree. The nonlinear threshold function f is performed at the root unit of each PE. The result, $x_j(1)$ then serves as the optical input to the output (Array 2) layer, which calculates

$$x_i(2) = f\{\textstyle\sum_j W_{ij}(2) \times x_j(1)\} \qquad [1b]$$

in a similar manner. A 16×4×1 scalable prototype[2] of this architecture has recently been constructed, with a preliminary performance measure in excess of 500,000 input presentations per second. An optimized performance measure of 10M presentations per second is expected.

Error backpropagation[4] is a gradient descent learning algorithm which minimizes a least-mean-squared error function defined on the weight state of the network. For brevity, we define this algorithm operationally on a three layer network. The *local unscaled error* of the hidden and output neurons, $h_j(1)$ and $h_i(2)$ respectively, are given by,

$$h_j(1) = \sum_i W_{ij}(2) \times e_i(2) \qquad [2a]$$
$$h_i(2) = y_i(2) - x_i(2) \qquad [2b]$$

where $y_i(2)$ is the desired output of the i-th neuron in the output layer and $x_i(2)$ is its actual output. These local errors become *scaled local errors* , ej(1) and $e_i(2)$, when multiplied by the derivative of the nonlinear function evaluated at the neuron's current activation,

$$e_j(1) = h_j(1) \times f'(a_j(1)) \qquad [3a]$$
$$e_i(2) = h_i(2) \times f'(a_i(2)) \qquad [3b]$$

The weights of the hidden and output neurons are adjusted according to,

$$\Delta W_{jk}(1) = \alpha \times e_j(1) \times x_k(0) \qquad [4a]$$
$$\Delta W_{ij}(2) = \alpha \times e_i(2) \times x_j(1) \qquad [4b]$$

where α is the learning rate. Unfortunately, the feedforward D-STOP neural system in Figure 1 is, as such, not well suited for implementing real-time error backpropagation learning. The large optical fan-out used to copy the input vector over each PE does not allow the necessary bidirectional communication and computational fan-in required by [2a]. Fundamentally, while the locations of the weights $W_{ij}(2)$ within Array 2 facilitate the parallel matrix-vector multiplication of [1b], the weights cannot be accessed in a manner which allows parallel execution of [2a]. which is also a matrix-vector multiplication using the same weight matrix, but in a transposed representation.

The tandem D-STOP architecture shown in Figure 2 can overcome these limitations. The arrays labeled "1" and "2" implement the feedforward mechanism of the model, as in Figure 2. The additional D-STOP array, labeled "2T", implements the transpose of the weight matrix for layer 2. That is, a second copy of the weight $W_{ij}(2)$ is stored at the i-th leaf unit of the j-th PE of this array. During the learning phase, the local scaled error $e_i(2)$ is calculated in the ith PE of array 2, according to [3b]. This term is distributed electrically to the leaf units within that PE, which, having also received the input signals optically from the previous layer, can update their weights according to [4b]. Simultaneously, each scaled error $e_i(2)$ is optically distributed to the i-th leaf of the every PE in Array 2T. In addition, the signal $x_j(1)$ from the previous layer is transmitted to the root unit of the j-th PE in array 2T. This signal is electrically distributed to every leaf unit within the PE. Two operations are performed in this array. First, the weights in array 2T must also be updated to maintain consistency with the weights of array 2. Since each leaf unit in this array receives $e_i(2)$ optically and $x_j(1)$ electrically, this update [4b] can be performed locally. Second, Array 2T computes the local, unscaled error terms [2a] for the previous layer. These results are obtained at the root units of Array 2T, then transmitted to the root units of Array 1. The local scaled errors $e_j(1)$ can then be calculated according to [3a] and electrically distributed to the leaf units, where the weights can be updated according to [4a]. In this manner, error backpropagation learning can be implemented on-line in parallel.

We will present the tandem D-STOP learning architecture in more detail, and discuss related systems issues. Methods of reducing hardware requirements and improving learning performance will also be discussed.

1. A. V. Krishnamoorthy, G. Yayla, and S. C. Esener, *IEEE Trans. Neural Networks*, 3, pp. 404-413 (1992)

2. G. Yayla, A. V. Krishnamoorthy, G. C. Marsden, J. Ford, V. Ozguz, C. Fan, S. Krishnakamur, J. Wang, and S. C. Esener, *SPIE 1992 Intl. Symp. on Optical Applied Science and Engineering*, San Diego, 1992

3. G. C. Marsden, A. V. Krishnamoorthy, S. C. Esener, and S. H. Lee, *Opt. Lett.* 16, pp. 1970-1972 (1991)

4. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, in Parallel Distributed Processing, vol. 1, MIT Press, 1986

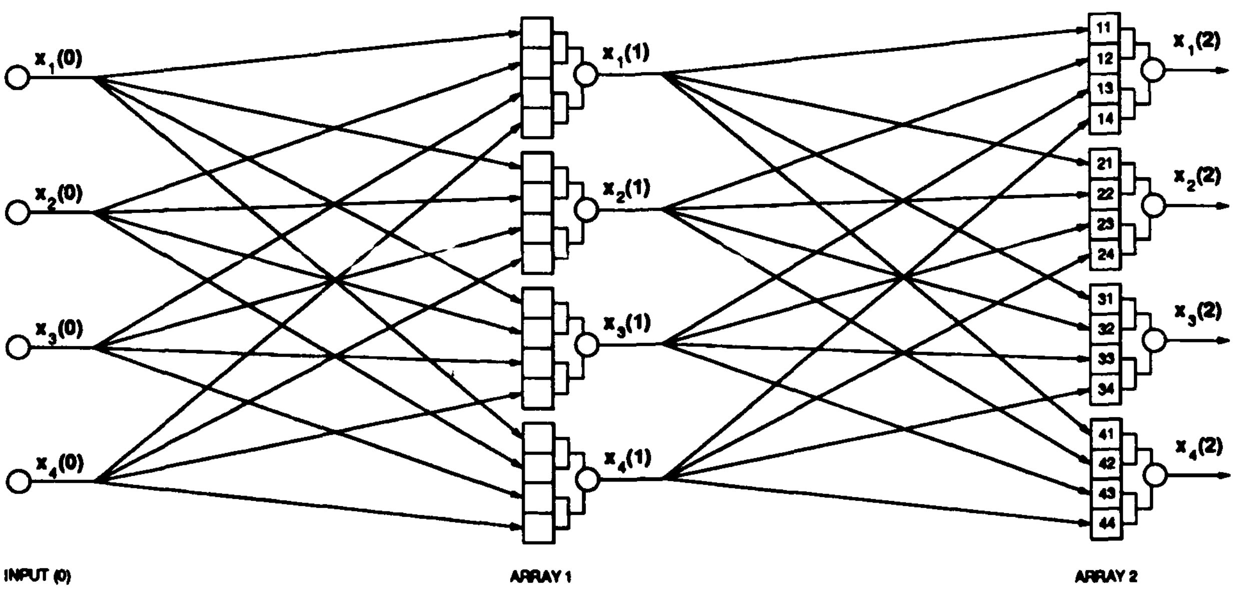


Figure 1: Cascaded D-STOP architecture for parallel recall.

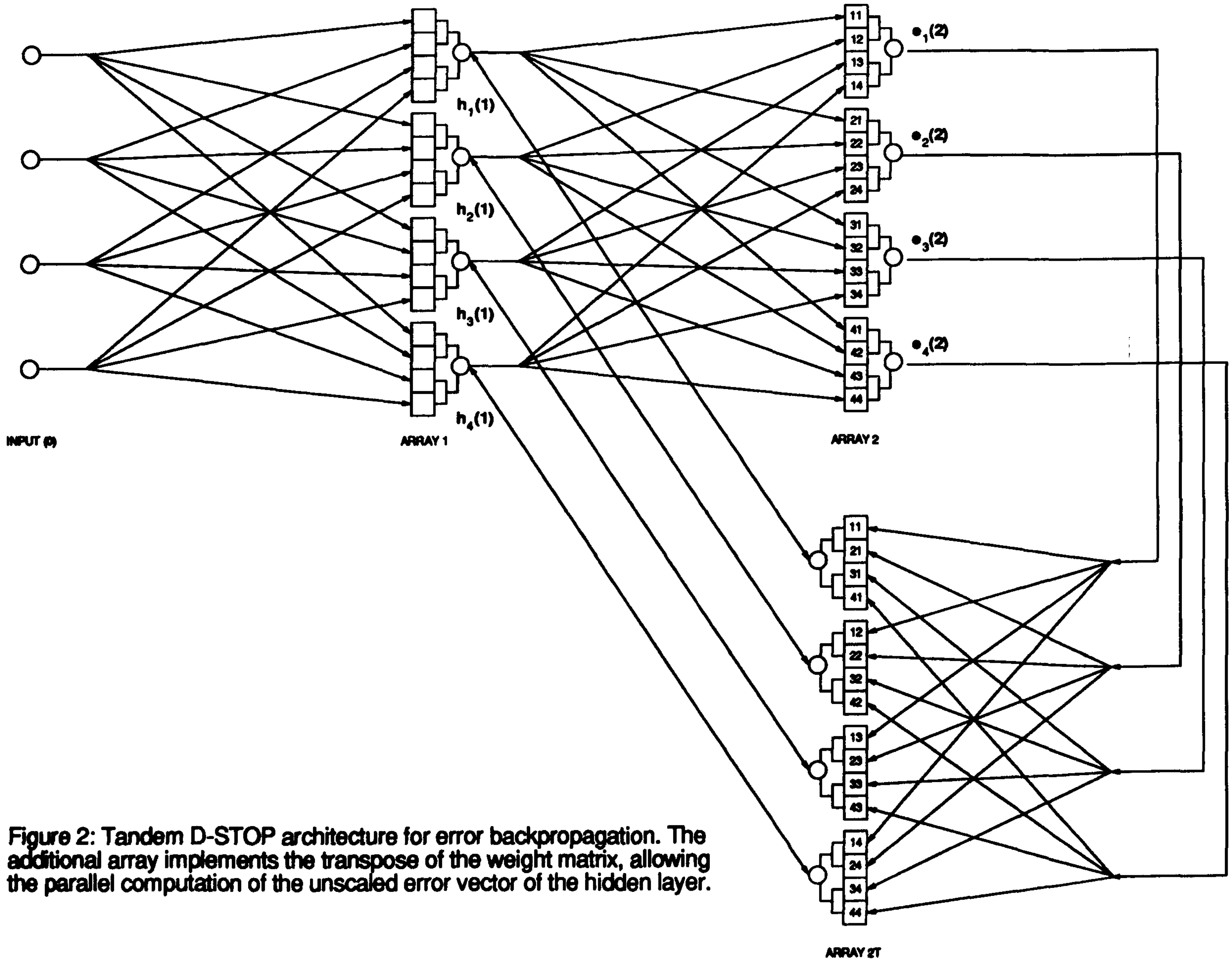


Figure 2: Tandem D-STOP architecture for error backpropagation. The additional array implements the transpose of the weight matrix, allowing the parallel computation of the unscaled error vector of the hidden layer.

Direct optical implementation of lateral inhibition and excitation.

Paul Horan.
Hitachi Dublin Laboratory, O'Reilly Institute, Trinity College, Dublin 2, Ireland.
Tel. +353-1-6798911 email: pghoran@hdl.ie

Introduction.

Networks utilising lateral inhibitory and excitatory interconnections have a number of interesting properties, particularly for unsupervised learning and self-organisation. Since the output of each node in a layer depends not only on the input from other layers but also on the state of many or all other nodes in that layer it is not possible to calculate the output of individual nodes separately but the solution for all nodes must be found in a parallel, self-consistent fashion. This is a problem both in simulation on conventional serial computers and in implementation, particularly in planar electronics. Optics, utilising the third dimension offers a possible solution.

Lateral interconnections can provide a number of different functions depending on the particular connectivity and the internal structure of the nodes. Large scale lateral inhibition can give rise to competitive or "winner-take-all" (WTA) behaviour within the field of inhibition.[1] In early vision processing the well known centre-on/surround-off or "Mexican hat" interaction provides contrast enhancement and forms the basis for much retinal processing.[2] The computational properties of laterally connected and recurrent networks have been extensively studied by Grossberg and co-workers over the years[3] and have evolved into the powerful class of ART networks.[4]

When inhibitory interaction is combined with local neighbour positive, excitatory connections *i.e.* centre/surround connections and Hebbian learning between the laterally connected layer and a prior, feed-forward layer a topographic or computational mapping can be established.[5, 6] Adaptation by Kohonen to aid calculation on conventional serial computers has allowed general exploration of these networks.[7] In tandem with theoretical studies there is extensive evidence from neuroscience as to the importance of topographic mappings.[8] As a definition of a topographic mapping Knudsen *et al* can be quoted; "In a computational map, there is a systematic variation in the value of the computed parameter across at least one dimension of the neural structure. .... they transform their input almost instantaneously into a place-coded probability distribution that represents values of the mapped parameter as locations of peaks in activity within the map".[8]

Lateral Inhibition.

To date, electronic and opto-electronic implementations of lateral interconnections have been inhibitory only and utilised planar electronic interconnection. Analogue VLSI implementations of "sensory" networks have been pioneered by Carver Mead and his group,[9] and a simple shunting network, with nearest neighbour inhibitory connections has been studied.[2] Winner-Take-All functionality with optical input and output has been achieved by two different mechanisms, both using a common global interconnection.[10,11] Both of these approaches are limited in the extent to which they can be scaled. Direct electrical interconnection is suitable for nearest neighbour and single global interconnection. The analogue VLSI "smart pixel" approach to signal processing within a node, as exemplified by the above references, offers the greatest possibilities for compact, robust systems. However, if we wish to implement a more complex or large scale network then there are extreme difficulties with planar, electrical interconnection. The combination of opto-electronic "smart pixels" with diffractive optics can address these problems as suggested in figure 1 for an inhibitory network.

Each pixel (as shown in the insert) consists of a photodetector, some processing electronics and a light source or modulator. Input can be optical or electrical. The light from any pixel is directed to its neighbouring pixels in a predetermined fashion by a diffractive optical element, the small fraction leaking through serving to provide the output. The signal from the photodetector provides the summed inhibitory signal from neighbouring nodes to update the pixel according to the internal electronics. This then becomes the output. The internal processing, whether additive or shunting[3] are not considered in detail here, only the means of interconnection.

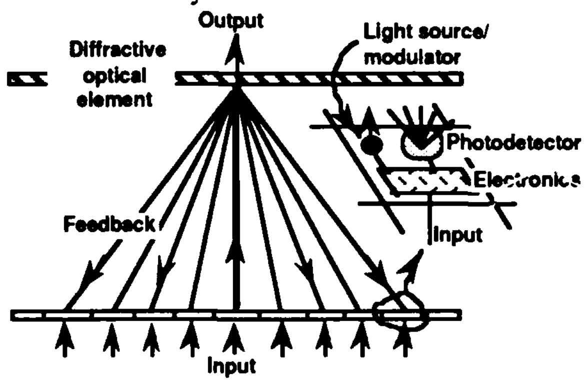


Figure 1 Optical implementation of lateral inhibition and (insert) pixel layout.

As with the electronic implementations described above, the extent of the lateral interaction is limited, in this case by the optical power available. If we wish to have winner-take-all behaviour it is necessary that the optical interconnection spans the distance between competing nodes. To ensure this we must invoke the principal of topographic organisation discussed above. This would imply that if a number of nodes are excited by the same input, then these nodes should represent somewhat similar properties. If the nodes are topographically organised then the excited nodes will be physically close, and so able to interact. Thus the network can be scale-up arbitrarily. Topographic organisation is imposed by the previous feed-forward layer, which we have not discussed here, rather than to the competing layer itself, but relies heavily on the type of processing described in the following section.

Lateral Excitation and Inhibition.

In the above section we considered only inhibitory connections. For certain applications, most notably the "mexican hat" centre/surround type interaction, both excitatory and inhibitory connections are required. This presents special problems for optical implementation, as light is a positive definite quantity. We shall consider this problem in the context of the centre/surround interconnection scheme. As shown in figure 2 the interaction (heavy line) can be broken down into a strong local excitation and a broader, weaker inhibition. This requires that from every node there are excitatory connections to near neighbours, and inhibitory connections extending to a larger radius. If an arbitrary input is presented to such a competitive layer stable "bubbles" of activity will form at the maxima of the input signal.[7] (For example see figure 4.)

This can be seen as WTA activity due to the lateral inhibition with the local excitation broadening out the winning node into a cluster. This type of behaviour provides great robustness and is common to biological processing. It is also essential for the formation of topographical mappings.[5] Figure 3 illustrates an architecture for optically implementing inhibitory and excitatory lateral interactions by spatially separating them in distinct layers. The node activity is calculated at Y_i, with the complementary X_i calculating the excitatory contributions and broadcasting the inhibitory signal. Thus a photodetector at each X_i sums the net excitatory input to that node. This is fed back to Y_i via a one-to-one electrical interconnection. The output of X_i is dependent on the net optical input and any external inhibitory inputs and is fed back to the Y_i according to the inhibitory interconnection strengths. Thus the photodetector at each Y_i calculates the net inhibitory input. The net activation of the node Y_i can be computed using the optically delivered inhibitory signal and the electrically delivered excitatory signal. This is then thresholded to give a new output. For suitable values of feedback parameters this network can form stable bubbles of activity, as illustrated in simulation in the following example. Such a network is at the heart of self-organising nets but it would still be necessary to have the inputs , at least partly, organised on a topographic basis to ensure possible winning nodes can compete.

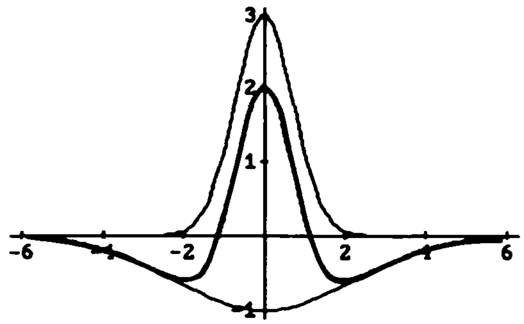


Figure 2 Centre/surround interaction (heavy line) and its decomposition.

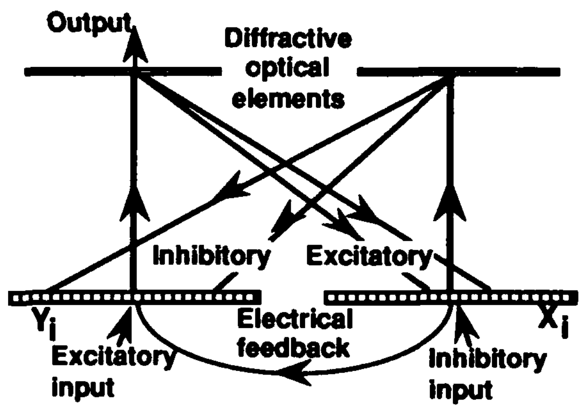


Figure 3 Lateral optical excitation and inhibition architecture.

We will consider a simple additive network with external excitatory inputs only. The network is initialised with $Y_{i(0)}$ = input pattern, $X_{i(0)} = 0$, and then allowed to iterate to a final state. The layers X_i and Y_i are alternatively updated according to equ's 1 and 2.

$$X_{i\,(t+1)} = X_{i(t)} + \alpha \sum_{-a}^{+a} Y_{i(t)} \qquad \text{Equ. 1}$$

$$Y_{i\,(t+1)} = \int_{th} \left(Y_{i(t)} + X_{i(t+1)} - \beta \sum_{-3a}^{+3a} X_{i(t)} \right) \qquad \text{Equ. 2}$$

where $\int_{th}$ is a threshold function. The values of Y_i after sequential iterations are shown in figure 4 along with the threshold function $\int_{th}$ used and the form of the lateral interaction. Values of feedback parameters $\alpha = .3$, $\beta = .04$, were used for input signal levels <.1 The interaction width a = 4, giving inhibition 3 times as wide as the excitation, and has been squared off to simplify calculations.

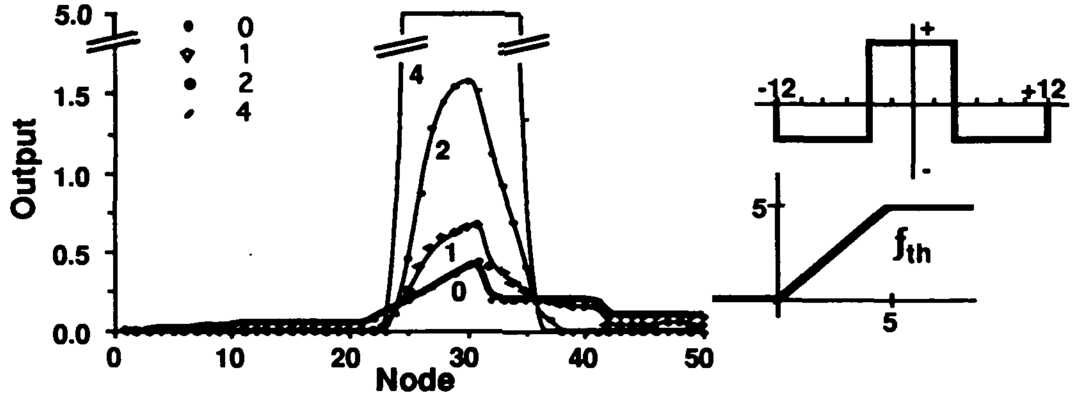


Figure 4 Simulation of the formation of a stable bubble of activity.
Inserts: lateral interaction strengths and the threshold function f_{th}.

As we can see a stable bubble of activity soon forms at the region of maximum input.

Conclusions.

If complex, hierarchical neural network systems are to be built then it will be necessary that individual sub-systems be to some extent unsupervised and self-organising. This will necessitate the use of lateral interconnections. The implementation of the dense and complex but regular and fixed connectivity required seems an area ideally suited to optics when combined with the possibilities of modern smart pixel technology.

References

1 Y. Cheng and Z Wan, "Distribution of winners in local inhibition", in *Proc. of the International Joint Conference on Neural Networks* (IJCNN 92) (IEEE 92CH3114-6, 1992) pp III 456-III 460.

2 B. Nabat and R. D. Pinter, *Sensory Neural Networks: Lateral Inhibition*, (CRC Press, 1991).

3 S. Grossberg, "Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors", Biological Cybernetics, 23, pp 121-134 (1976).

4 G. A. Carpenter and S. Grossberg, "ART 2: self-organisation of stable category recognition codes for analog input patterns", Applied Optics, 26, pp 4919-4930 (1987).

5 C. von der Malsburg, "Self-organisation of orientation sensitive cells in the striata cortex", Kybernetik 14, pp 85-100 (1973).

6 A. Takeuchi and S. Amari, "Formation of topographic maps and columnar micro structures in nerve fields", Biol. Cybernetics, 35, pp 63-72 (1979).

7 T. Kohonen, *Self-Organisation and Associative Memory*, (Springer-Verlag, 1984).

8 E. I. Knudsen, S . du Lac, and S. D. Esterly, "Computational maps in the brain", Annual Review of Neuroscience, 10, pp 41-65 (1987).

9 C. Mead, "*Analog VLSI and neural systems*", (Addison-Wesley, 1989).

10 C. V. Radehaus, J. I. Pankove. M. Kuijk, P. Hermans, and G. Borghs, "Maximum detection with a two-dimensional optoelectronic p-n-p-n winner-take-all network", Applied Optics 31, pp 6303-6306 (1992).

11 K. Wagner and T. Slagle, "Competitive optical learning with winner-take-all modulators", in *Optical Computing*, Vol. 6 of 1991 OSA Technical Digest Series (OSA, 1991) paper WA5.

Neural Algorithms that Exploit Optical Hardware Characteristics

John F. Snowdon
Department of Physics, Heriot-Watt University
Edinburgh EH14 4AS, U.K.

Since the beginning of the respective subjects of neural nets and optical processing, groups in both disciplines have believed that the fusion of neural processing ideas with optical technology will produce the most powerful computers of the future. The interconnection density and bandwidth available in optics is a natural platform on which to conceive neural processors[1] but a lack of devices has held up progress. Most approaches at present attempt to implement known neural models and algorithms in optical hardware. An alternative approach may be to invent new, or adapt old, models and algorithms to map onto the optical hardware that exists. In particular the dynamics of devices play an important role in asynchronous networks, and these dynamics may be exploited. This paper presents applications of nets in which some attempt is made to exploit these ideas. Amongst the possibilities raised by such an approach are exciting new 'neural' models and a synthesis of digital and neural techniques.

The devices considered here are Non-Linear Interference Filters (NLIFs) and S-SEEDs. These devices are chosen primarily because of the extensive experience that has been gained in the practical use and theoretical study of such devices by the optical computing group at Heriot-Watt University. The devices may be exploited for network use in many ways - different aspects of their characteristics may be used for weights, thresholds or signum functions. Many different characteristics are available from this class of devices depending on how (and if) they are clocked, on the method of supplying their operating power, and on the quantity of power supplied.

Further sets of functions may be generated by dynamical means and by coupling devices in specific ways.

The simplest conceptual form of a neural network model is a weighted crossbar interconnect or matrix vector multiplier. The neurons are a vector of thresholding pixels whose outputs are fanned-out and passed through a weighting matrix and then fanned-in in a direction orthogonal to the initial fan-out to form the off-axis inputs of the thresholding devices. Both Hebbian and Error propagation schemes[2] may be mapped onto hardware of this form. The weights may be adjusted using electronic hardware coupled to an electrically addressed SLM, but the goal is to produce simple enough algorithms so that optical learning feedback may be employed. This is most desirable since the feedback process is likely to be computationally the most expensive.

One application which we consider is the replacement of various functional units in an optical cellular logic image processor[3,4,5], in order to build in fault tolerance. As an example consider an adaptive programmable logic unit (APLU) using error propagation. The APLU must be able to learn any desired logic function from multiple presentations of the training set and must be able to re-adapt to restore that functionality in the event of unit failure.

An example of the dynamic evolution of the learning process is shown in Fig. 1(a) and 1(b). The task is delineated into two phases. In the first phase, the logic unit learns a simple task, the solution of the desired logical operation (in this case the Ex-Or problem) starting from a random weight matrix. The neuron states evolve according to the following equations[6],

$$\frac{\partial\theta_i}{\partial t} = b(\theta_i)\,P_{Hi} - 1 - \frac{\theta_i}{c} + \sum_{j=1}^{N} W_{ij}\,(1 - ab(\theta_j))$$

and

$$P_{0i} = (1 - ab(\theta_i))\,P_{Hi}$$

In this example, only three of the neurons evolve since neurons 1 and 2 are taken to be inputs, or input units, and hence $\dot{\theta}_{1,2} = 0$ and their outputs set to the desired steady state characteristic values. The third neuron is declared the output unit and its output along with the target signal forms the learning signal for the weight update. The other three neurons effectively constitute hidden units in that they form part of an asynchronous nonlinear link between the inputs and the outputs, but are not explicitly tampered with. The selection of 3 units to perform the task is accidental, the network chooses to use only three neurons in a wide variety of the initial random weight matrices used in the experiment and, indeed, the weight matrices generally converge to similar numerical values.

Error may be propagated in a wide variety of ways, each with different dynamics and features. The primary assumption made in this experiment is that the weight updates take place on a larger timescale than the network evolution so that some of the 'decision making capacity' of the system is being utilised during update. Indeed, it may be desirable to clock the weight update system so that the inputs are sampled at well defined intervals. Precise trade-offs of learning efficacy against update speed are difficult to establish at present. The speed and nature of the algorithm employed varies with the type of hardware envisaged. In the experiment shown in Fig. 1(a) correct operation generally ensues over less than 10 presentations of the input set. The particular learning algorithm employed in this experiment is to update the weights according to

$$\frac{\partial W_{ji}}{\partial t} = \eta g((t_z - P_{0z})\,P_{0i}) \qquad j \neq z$$

and

$$\frac{\partial W_{ji}}{\partial t} = \eta g\,((P_{0z} - t_z)\,P_{0i}) \qquad j = z$$

subject to $0 < W_{ji} < \beta$ and $W_{ii} = 0$, where η represents the time constant of the learning/weighting hardware and g is the transfer function of the hardware. The subscript z designates the unit(s) selected to be outputs, and t_z represents the target for that output on a particular learning presentation. β fixes the dynamic range of the weights (which are all positive).

The second phase of the experiment involves failing one of the hidden units, in this case, unit 5 fails low. The system relearns to perform its task correctly in less than 100 presentations, but finds this relearning generally more difficult (due to the effective spurious input). In some cases the system will begin to utilise other hidden units, in some cases it manages to relearn the task on an effectively smaller set of units. Of course, failure of the output unit is catastrophic to the process, and must be dealt with by either incorporating additional redundancy or by reallocating the output status to a hidden unit in the event of an unreasonable number of failures.

In conclusion, a neural system using optics or optoelectronics has been presented in a novel application. The employment of real device dynamics in the system and in the algorithms, appears to give rise to highly successful learning and convergence rates in the simulations. Tables of these rates (based on experimental device characteristics) over a wide range of tasks (including pattern recognition and content addressable memory) will be presented at the conference.

References
1. Y.S. Abu-Mostafa and D. Psaltis, "Optical neural computers", Sci. American, March 1987.
2. D.E. Runelhart, G.E. Hinton and R.J. Williams, "Learning internal representations by error propagation", "Parallel Distributed Processing: Explorations in the Microstructure of Cognition", ed. Runelhart and McClelland, Bradford Books, Cambridge, Mass., 1986.
3. J.F. Snowdon, A. Kashko and B.S. Wherrett, "Design of digital optical processors", SPIE, 1772, 1992.
4. B.S. Wherrett et al., "Construction and tolerancing of an optical CLIP", SPIE, 1215, pp. 264-279, 1990.
5. S. Wakelin, G.R. Smith and F.A.P. Tooley, "Design and construction of a cellular logic image processor using symmetric SEED arrays", Proc. OC'94, Minsk, 1992.
6. B.S. Wherrett and J.F. Snowdon, "Tolerance studies for digital optical computing", Int. J. Opt. Comp., 1, pp. 41-70, 1990.

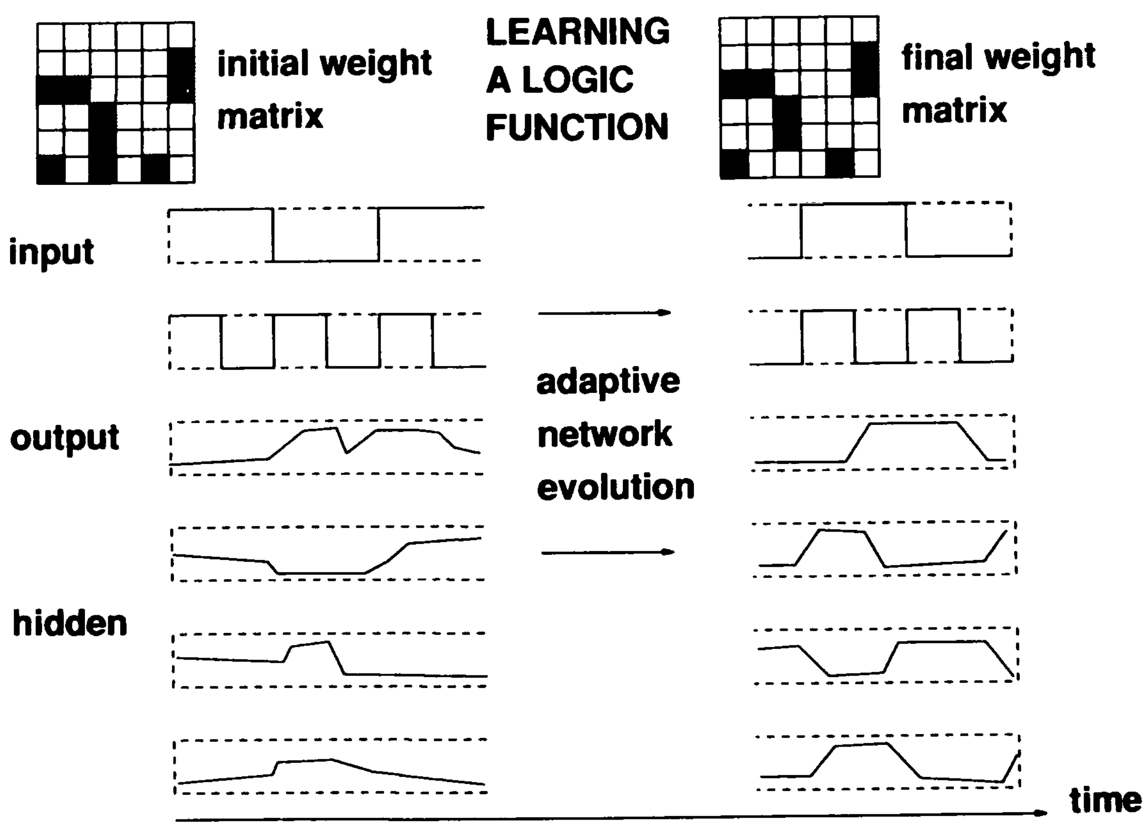


Fig. 1(a). The NLIF pixel outputs and weight matrix as the network learns a logic function.

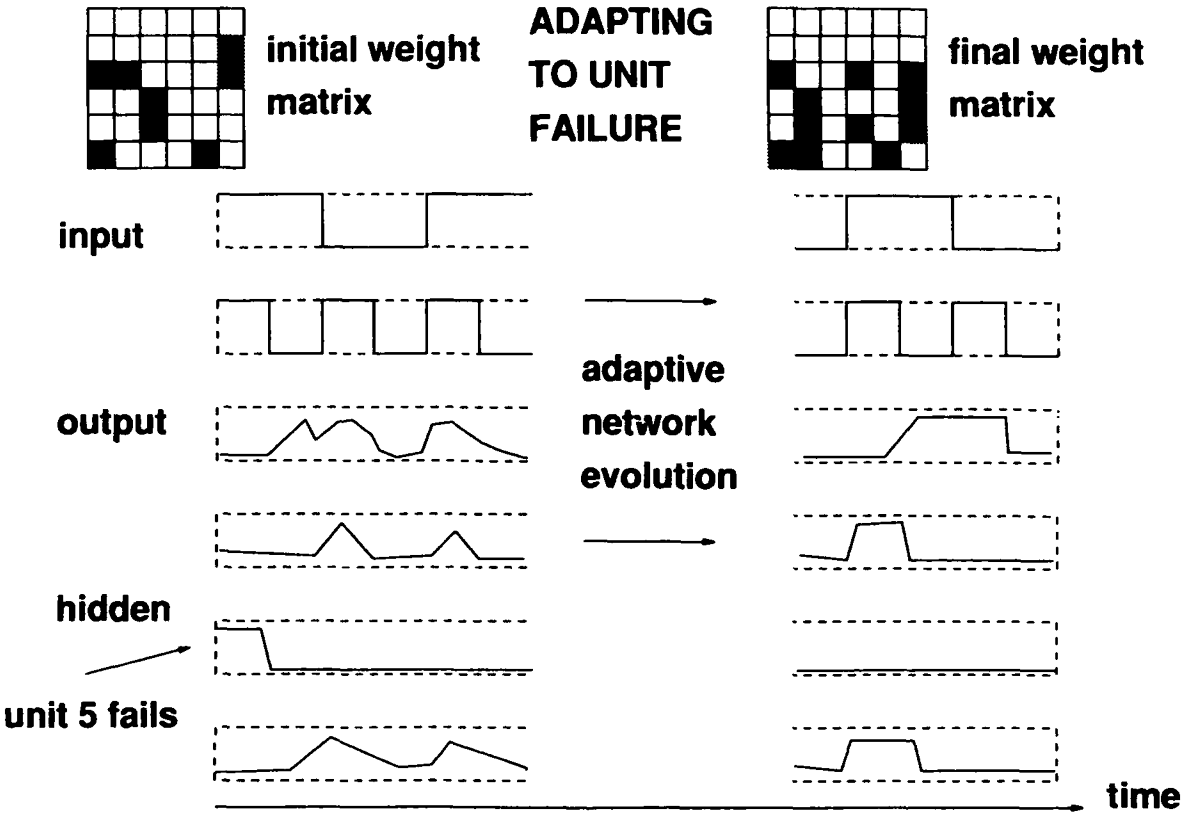


Fig. 1(b). The NLIF pixel outputs and weight matrix as the network re-learns a logic function in the event of a unit failure.

OPTICAL PROCESSING SYSTEM FOR INDUSTRIAL PROCESS CONTROL APPLICATION

Mrs. Sushila Singh and A.D.Shaligram

Department of Electronic Science, University of Poona,
Pune 411 007, India.

1. INTRODUCTION :

Process control is a time based concept.The objective is to maintain a property at some specified value in time as the manufacturing process proceeds. For most of the process control applications , real time control is very much desired. Thus the speed of the control system is of prime importance. In this paper a novel concept of use of a optical computing block for the real time control application is attempted. Optical computing inherently has the advantages of high speed and massive parallelization.

The concept, design, fabrication details and testing of an optical magnitude comparator are described. An optical equivalent of TTL 7485, magnitude comparator has been designed and fabricated. The design of optical gates is based on shadow casting method [1]. For practical implementation of the system an LED matrix and Liquid Crystal based SLMs and masks are used. The fabricated comparator, after testing its perfoemance, is further used in a temperature controller.

2. PROCESS CONTROL BASICS :

The following example shows the strategy used for regulation or control of a variable. Equation (1) represents the state of a process as a function of process variables and time [2].

$$\nu_j = G\,(\nu_1, \nu_2, \nu_3, \ldots\ \nu_m, t) \qquad \ldots \qquad (1)$$

where,

ν_j : the state variable
G : the functional representation of the process
ν_m: the process variables
t : time.

The variable ν_j, depends on other variables in the process and also on time. This can be called as **controlled variable**. Select one variable, say ν_1, in equation 1 to be a **controlling variable**. Thus, ν_1 , will be a variable that affects the value of ν_j . For temperature control problem this could be the heater current affecting the temperature.

Make a measurement of controlled variable ν_j, to determine its present value. Compare the measured value of controlled variable with desired value for maintenance of the product property. The desired value is called **set point** of the controlled variable. Determine a change in the controlling variable that will correct any deviation or error of the controlled variable from the set point. Feedback this changed value of controlling variable to the process to create a correction to the controlled variable. If this procedure is repeated on a regular basis , the result is regulation of variable in time. Furthermore, if such a process is set up for all the variables on which all the product properties depend, the result is called process control. The role of electronic comparators for indicating the difference (**error**) between measured value and set value, in traditional process control is well known . This error is then processed further to reduce it to minimum value. Fig. 1 shows the block diagram of a process control loop.

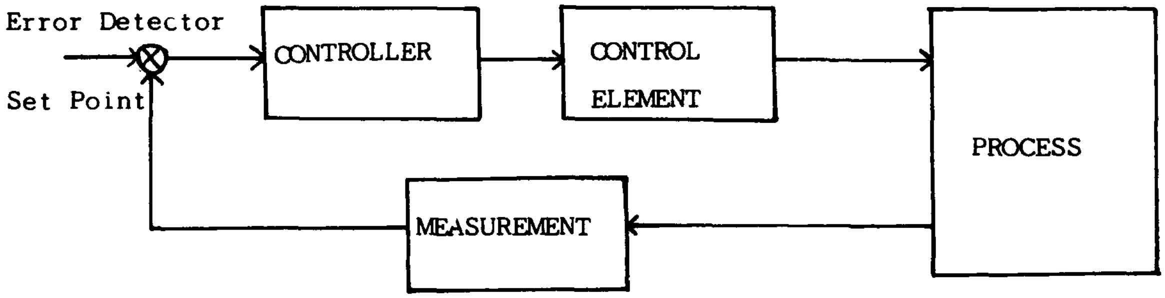


Fig.1 Blockdiagram of a process control loop.

In the present work an attempt has been made to implement the comparator optically. The optical implementation has advantages of parallelism and high speed. Therefore if large number of parameters are to be compared then optically implemented comparator would be very much useful.

3. DESIGN OF THE OPTICAL MAGNITUDE COMPARATOR :

The process contol can be achieved both by analog as well as digital techniques. There are advantages and limitations of both the approaches. However, a digital control is accpeted widely as a stable control. The controlled variable is typically in analog form, which is degitized with a suitable ADC and the digital data is further processed for the control operation. A digital comparator would thus be a block finding the error between the set and the sensed values of controlled variable. NE7485 is a standard TTL magnitude comparator[3] which compares two 4-bit binary numbers. The attempt has been made to develop a optical magnitude comparator. This would form one block of the optical processor developed for temperature control application.

The implementation of optical magnitude comparator makes use of shadow casting method given by Tanida et al . Using this technique **sixteen** different logic functions can be implemented. Based on these logic functions we have designed the **Optical Magnitude comparator.**

4. Optical Magnitude Comparator Using Shadowcasting Approach :

The magnitude comparator is of much help because it can compare a binary word with another. Here advantage is that the operation of comparison is made simultaneously on all bits in the word. Here we have reported a design of **OPTICAL MAGNITUDE**

COMPARATOR based on above mentioned technique. Fig.2 shows the block diagram of the same.

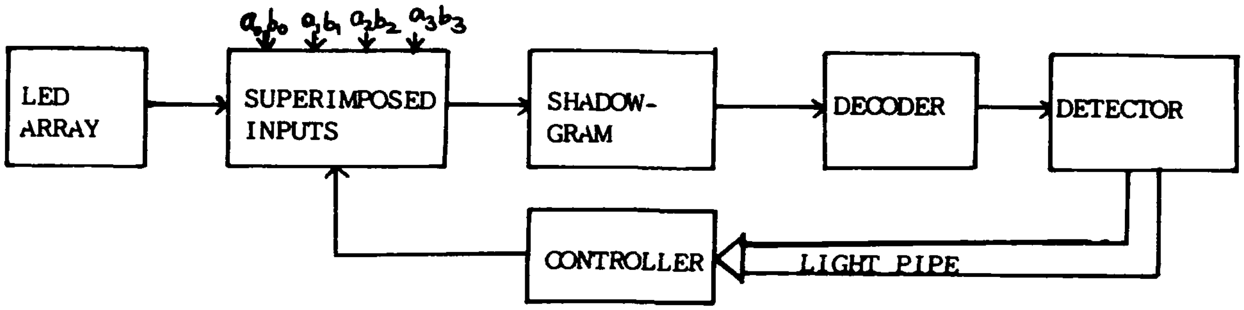


Fig.2 Blockdiagram of Optical Magnitude Comparator

We have designed it for comparing two four bit binary numbers. Plane P1 consists of light sources, mainly LEDs. Their on-off conditions can be generated by using a controller block. For comparison of four bit word we require source array of size 12 X 12. The word A consists of four bits labelled as A3, A2, A1 and A0. The word B also consists of four bits labelled as B3, B2, B1 and B0. Both A3 and B3 are most significant bits. Each bit in the word is spatially encoded [1]. While comparing two input binary objects their spatially encoded forms were superimposed on each each other by keeping the one object in front of the other. Using shadowcasting method a gate level form of TTL NE7485(electronic magnitude comparator) was implemented optically.

In the on-off type of temperature controller this optical magnitude comparator was inserted in place of electronic digital comparator to compare the difference between set and sensed temperature. The blockdiagram of the optical processor used for temperature control application is shown in fig. 3. Performance of the temperature control system with conventional electronic and the optical magnitude comparators is studied. Results are quite encouraging.

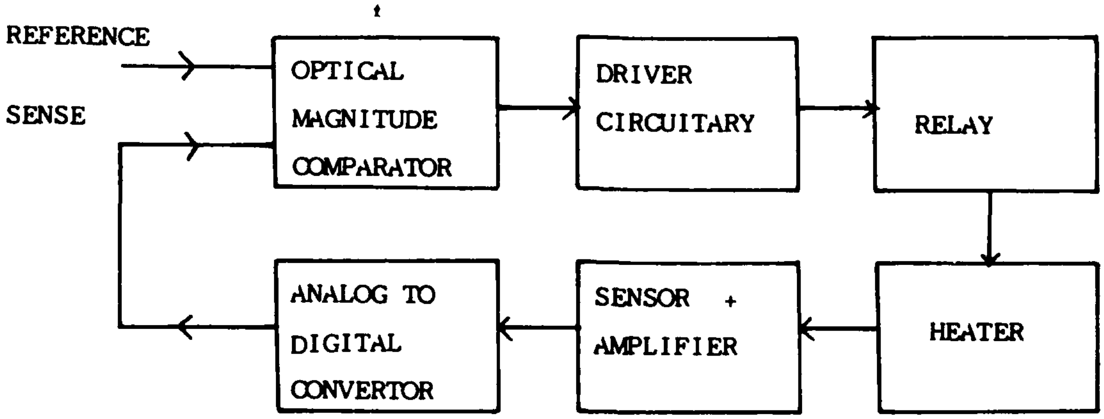


Fig. 3 Blockdiagram of optical processor for temperature control application.(ON-OFF)

This optical processor is also used in feedforward type of temperature controller and its performance is tested with conventional feedforward type of temperature controller.

5. REFERENCES :

[1] Y. Ichioka and J. Tanida; "Optical Parallel Logic Gates Using Shadow Casting Sysytem for Optical Digital Computing"; Proceedings of IEEE, **Vol 72**(4), 787-801 (1984).

[2] C.D. Jonson, "Microprocessor Based Process Control", Prentice-Hall Inc. (1984), pp 3.

[3] 7485 : "54/74 Series 4 bit Magnitude Comparator"; Signetics TTL Data Manual, pp 128-131.

AN OPTICAL ALGORITHM FOR A PLANE IMAGE RECONSTRUCTION FROM ITS BOUNDARY

Y. B. KARASIK
Computer Science Department, School of Mathematical Sciences,
Tel Aviv University, Tel Aviv, Israel
e-mail: karasik@math.tau.ac.il

1 Introduction

As is known optical analog processors can readily extract boundary ∂A from any plane image A by means of both optical differentiation of the image [10] and optical symbolic substitution (SS) [5] [3].

However, the inverse operation – reconstructing a plane image from its boundary usually called region filling, cannot be performed optically so easily. Despite this operation having many applications in scene analysis, pattern recognition and navigation of mobile robots, it has received no attention from workers in optical image processing and no approach has been suggested for its optical realization.

Meanwhile, this problem has received much attention from computer scientists who have proposed many solutions intended for ordinary display devices [8], [7], the best of which [1], as far as we know, has $O(n \log N)$ time complexity, where n is the number of pixels contained in the boundary and N is the number of pixels contained in the diameter of a polygon.

Due to their high time complexity, these algorithms are poor on real time image reconstruction which is required in many applications (e.g. mobile robot navigation [6]). Hence, the problem emerges of applying optical image processing techniques to developing a new filling algorithm which takes into account that many complicated operations on images can be performed optically in constant time and, hence, can be used as new primatives in developing an optical filling algorithm capable of filling the interior of a region hopefully also in constant time. We keep in mind the following operations:

• **addition/subtraction:** $C(n,m) = A(n,m) \pm B(n,m)$ of two images in parallel for each pixel (n,m) [2].

• **scalar multiplication**: $C(n,m) = k \cdot B(n,m)$ which is obtained as a result of the passage of light through both the SLM containing $B(n,m)$ and the SLM whose pixels have optical density k. We assume that photodetectors cannot distinguish between two electromagnetic waves if their amplitudes differ by less than unit under appropriate scaling. Thus, elements of the matrices we deal with are positive integers and scalar multiplication of an image is, in fact, the operation: $C(n,m) = \lfloor k \cdot B(n,m) \rfloor$.

• **thresholding of an image at a given level L** [4]:

$$C(n,m) = Threshold_L(B(n,m)) = \begin{cases} 1, & \text{if } B(n,m) \geq L; \\ 0, & \text{otherwise;} \end{cases}$$

• **convolution** [9]:

$$C(n,m) = \sum_{i=1}^{N} \sum_{j=1}^{M} A(n-i, m-j) \cdot B(i,j) = A(n,m) * B(n,m);$$

Based on these operations as primatives, we can try to reduce time complexity of the filling and propose constant time filling algorithm.

2 Description of the algorithm

To obtain the idea of the algorithm turn from a discrete plane consisting of pixels to a continuous plane consisting of points and consider any horizontal line $y = a$ intersecting the plane figure which is assumed to be a polygon (see fig. 1).

What segments of this line lie in A? To answer the question we should assign a weight to each intersection point $(x_i(a), a)$ by the following recurrence formulae:

$$weight(x_i(a), a) = \sum_{k=1}^{i-1} weight(x_k(a), a) + \begin{cases} 1, & \text{if the } i^{th} \text{ intersection is of} \\ & \text{the first type (see fig. 2 a, b);} \\ \frac{1}{2}, & \text{if the } i^{th} \text{ intersection is of} \\ & \text{the second type (see fig. 2 c, d);} \\ -\frac{1}{2}, & \text{if the } i^{th} \text{ intersection is of} \\ & \text{the third type (see fig. 2 e, f);} \\ 0, & \text{if the } i^{th} \text{ intersection is of} \\ & \text{the fourth type (see fig. 2 g, h),} \end{cases}$$

where i is the number of intersection in order from left to right.

It is easy to see that each segment of the line whose left endpoint has an odd weight, lies in A.

We can also attribute a weight to other points (x, y) by the following formulae:

$$weight(x, y) = \sum_{x_i(y)<x} weight(x_i(y), y).$$

Obviously, only points with an odd weight lie in A.

Now attribute a type to each point of ∂A as follows: the type of a boundary point is the type of the intersection between ∂A and the horizontal line passing through this point. Since interior points of horizontal boundary segments cannot be intersection points, we attribute to them the fourth type. Thus, we have:

$$\partial A = \bigcup_{i=1}^{4} S_i,$$

where S_i is the set of boundary points of the i^{th} type.

Theorem 2.1

$$weight(x, y) = S_1 * ray(x, y) + S_2 * \frac{1}{2} \cdot ray(x, y) + S_3 * -(\frac{1}{2}) \cdot ray(x, y),$$

where

$$ray(x, y) = \begin{cases} 1, & if\ x \geq 0, y = 0; \\ 0, & otherwise, \end{cases}$$

and convolution between set S and any function $f(x, y)$ is defined as follows:

$$S * f(x, y) = \sum_{(p,q)\in S} f(x - p, y - q).$$

We omit the proof in this version.

Let us now turn from the plane consisting of points to the plane consisting of pixels and generalize the approach stated above.

Instead of the function $ray(x, y)$ consider the image

$$RAY(n,m) = \begin{cases} 1, & \text{if } m = 1, n \geq 1; \\ 0, & \text{otherwise.} \end{cases}$$

Instead of boundary points of the i^{th} type consider boundary pixels of the i^{th} type (see fig. 3). We assume here that each boundary pixel is adjacent to exactly two others because those boundary pixels which do not satisfy this property can be removed from the boundary without breaking its topology with the help of SS (we omit details in this version).

Let S_i be now a set of boundary pixels of the i^{th} type and

$$WEIGHT(n,m) = S_1 * 2 \cdot RAY(n,m) + S_2 * RAY(n,m) - S_3 * RAY(n,m).$$

It is easy to see that the interior of A consists only of those pixels (n, m) which have $WEIGHT(n, m)$ not a multiple of 4.

Hence, we can propose the following filling algorithm.

Step 1. Remove redundant corner pixels from the boundary. As a result we obtain new boundary (of the same region) where all pixels are of type 1, 2, 3 or 4.

Step 2. For each $i = \overline{1,3}$ recognize boundary pixels of the i^{th} type . For this aim correlate the boundary with masks depicted on fig. 3 respectively and threshold the resulting images at level 3. As a result we obtain images S_1, S_2 and S_3.

Step 3. Compute

$$WEIGHT(n,m) = S_1 * 2 \cdot RAY(n,m) + S_2 * RAY(n,m) - S_3 * RAY(n,m).$$

Step 4. Compute

$$W(n,m) = \lfloor WEIGHT(n,m)/4 \rfloor.$$

Step 5. Compute

$$INTERIOR(n,m) = WEIGHT(n,m) - 4 \cdot W(n,m) = \begin{cases} 2, & \text{if the pixel } (n,m) \text{ is interior;} \\ 0, & \text{otherwise.} \end{cases}$$

Since each step of the algorithm requires constant time for its performance, we obtain:

Theorem 2.2 *The interior of a plane image can be filled optically in constant time.*

References

[1] H. H. Atkinson, I. Gargantini, and T. R. S. Walsh, Filling by quadrants or octants, *Computer vision, graphics, and image processing*, vol. 33, pp. 138–155 (1986).

[2] H. Fujiware, K. Nakagawa, Real-time image subtraction and addition using two cross-polarized phase-conjugate waves, *Optics Communication,* 79(1, 2), 6-10, 1990.

[3] S. D. Goodman and W. T. Rhodes, Symbolic substitution applications to image processing, *Applied Optics,* 27, pp. 1708–1714 (1988).

[4] C. Gu, S. Campbell, J. Hong, Q. B. He, D. Zhang, P. Yeh, Optical thresholding and maximum operations, *Applied Optics,* 31(26), pp. 5661–5665 (1992).

[5] A. Huang, Parallel Algorithms for Optical Digital Computers, *Proc. Tenth International Optical Computing Conference* (IEEE Computer Society, Los Angeles, 1983), pp. 13–17.

[6] Y. Y. Huang, Z. Cao and E. L. Hall, Region filling for mobile robot using computer graphics, *Proc. of 1986 IEEE International Conference on Robotics and Automation*, pp. 1607–1614 (1986).

[7] T. Pavlidis, Filling algorithms for raster graphics, *Computer graphics and image processing*, vol. 10, pp. 126–141 (1979).

[8] U. Shani, Filling regions in binary raster images: a graph-theoretical approach, *Computer graphics*, vol. 14, pp. 321–327 (1980).

[9] A. B. VanderLugt, Signal detection by complex spatial filtering, *IEEE transactions, Information Theory*, IT-10(2), April 1964, pp. 139–145.

[10] S. K. Yao and S. H. Lee, "Spatial differentiation and integration by coherent optical correlation method", Journal of Optical Society of America **61**, 474-477 (1971).

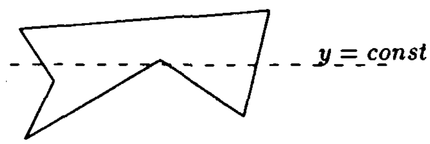


Figure 1: a polygon being intersected by a horizontal line.

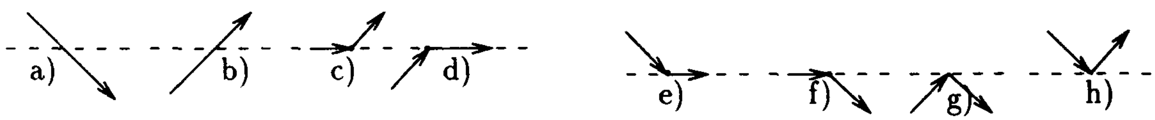


Figure 2: four types of intersections; directions of boundary segments are shown by arrows, points of intersections are indicated as rings.

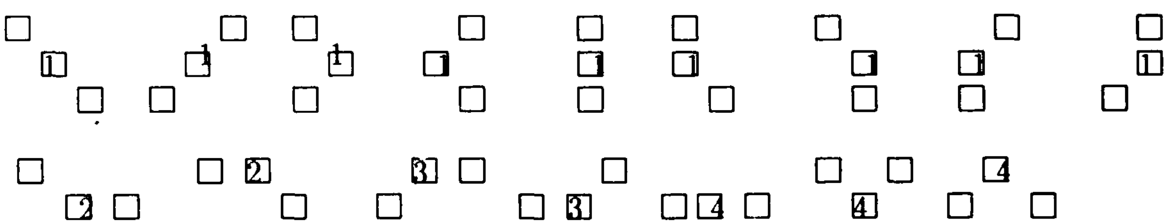


Figure 3: boundary pixels of the first, second, third and fourth types marked as 1, 2, 3 and 4 respectively.

OPTICAL ALGORITHMS IN GEOMETRY

Y. B. KARASIK
School of Mathematical Sciences
Tel Aviv University
Tel Aviv 69978, ISRAEL

M. SHARIR
School of Mathematical Sciences
Tel Aviv University
Tel Aviv 69978, ISRAEL, and
Courant Institute of Mathematical Sciences
New York University
New York, NY 10012, USA

Abstract

We consider the application of optical computational operations to problems that arise in various geometric applications, such as computer graphics, image processing, pattern matching, etc. We show how to implement, using only a constant number of basic optical operations, a variety of algorithms that solve fairly complex tasks of this sort. We discuss the methodology that we propose, list the geometric algorithms that we have developed to date, and illustrate the approach in two more detailed applications, motivated by computer graphics and pattern matching: drawing all possible lines connecting pairs in a given set of points, and computing the closest pair (or k closest pairs) of points in a given set.

In the last several years there has been significant progress in the design of optical devices that can perform computational tasks. The input to such devices is a *spatial light modulator* (*SLM* for short), which we can regard as representing a bivariate function $f(x, y)$ encoded in some optical form, such as a plane light wave with variable amplitude, or with variable phase, or as a planar slab made of optical material with variable transmittance coefficient, etc. The optical device then takes the input SLM's, and pass them through various lenses, mirrors, filters, holograms, possibly also interacting several input SLM's with each other. In this manner a variety of output SLM's can be generated, some of which are fairly complex functions of the input SLM's. In particular, in constant time, one can compute various pointwise arithmetic operations on the input SLM's [6], compute Fourier transforms [8], convolutions [10], conformal transformations of the coordinate system [9], thresholding the image at a certain level of brightness [5], and many other operations. The output SLM's can then be stored in various optical manners, and then be used as inputs for further optical operations.

These technological advances open up the possibility of constructing a full-fledged general-purpose optical computer, with a reasonably large storage of optical SLM's, and with a processing unit capable of executing any program consisting of optical and other auxiliary operations. At present such computers are not yet fully available, although various projects are currently studying the construction of prototypes of such computers. In spite of some technical difficulties that still exist, one can anticipate that such devices will become available in the not-too-distant future.

From an *algorithmic* point of view, optical computing opens up immensely powerful new opportunities for speeding up the solutions of many basic problems that have been studied by computer scientists using conventional computers. This is because various complex operations, that normally are fairly expensive, can now be performed in constant time. Thus we propose to build upon the standard library of optical operations a variety of composite operations (which we call *optical algorithms*), that perform progressively more complex tasks. Besides the theoretical interest in such an approach, we believe that it will be beneficial to the design of future optical computers, because it highlights desired design issues, whose implementation will make those computers a much more convenient tool. It is our hope to make engineers, who are currently working on the design of optical computers, aware of these 'software' issues, so that future theoretical research on optical algorithms might influence their work. There have been a few recent studies of a similar flavor; see e.g. Reif and Tyagi [7] (see also [1], [2] for additional related works).

Geometric algorithms form an ideal area for the development of optical algorithms, because the basic optical 'data type', the SLM, is an ideal medium for representing (two-dimensional) geometric data. In fact, in many geometric algorithms (in robotics, computer vision, image processing, object recognition, etc.) the data already exist in the form of 2-d images, so direct optical manipulations of such images, besides being ultra-fast, will also save the need to transform such images from raster to vector representation and back. In more 'artificial' cases, when the input to a problem is given in a discrete combinatorial fashion (e.g. a set of n points or n lines in the plane), there arises the technical issue of creating an SLM containing this input. This can be done using fairly standard techniques (see a previous work by the authors [11] for more details), and in this paper we will ignore this issue. Once the input SLM has been created, we can apply to it the whole arsenal of optical operations and obtain very fast algorithms.

Building upon the basic optical operations mentioned above, we first note that they facilitate the implementation of many basic geometric operations in a constant number of optical steps. Some of these operations are (see [11] for details of the implementation of these operations):

(i) Union/intersection of plane figures (here the input SLM's are assumed to store the characteristic functions of the given figures);

(ii) The Minkowski sum/difference of plane figures (same form of input SLM's as above):

$$A \pm B = \{a \pm b \mid a \in A,\ b \in B\};$$

(iii) The Radon (or Hough) transform [4], which serves as a basis for optical realization of *geometric duality*, which maps in the plane points (a, b) to lines $ax + by = 1$ and lines to points via the inverse of this mapping (see [3] for more details).

Using these building blocks, we have developed efficient optical algorithms for many basic geometric problems, including the following:

- Perform a variety of geometric duality transformations.
- Compute the *convex hull* of a planar point set.

- Compute the *Voronoi diagram* and its dual *Delaunay triangulation* of a planar point set.
- Compute the portion of a polygonal region *visible* from a given point or edge.
- Perform multiple *ray shootings* in a given planar region.
- Solve *linear programming* problems in the plane.

More details, and a more comprehensive list of our geometric algorithms, are given in [11, 12].

In this abstract, we exemplify this methodology by two simple optical algorithms that solve rather non-trivial problems.

Drawing Straight Lines Between all Pairs of Given Points: Our first problem is motivated by computer graphics. We are given a point set $P = \{p_i = (x_i, y_i)\}_{i=1}^{N}$ on an optical SLM, and we wish to draw, on an output SLM, all possible straight lines passing through each pair of points in P. We can solve this problem as follows:

> Step 1. Dualize the points $\{p_i\}_{i=1}^{N}$ into the respective straight lines $\{(x, y)|x_i \cdot x + y_i \cdot y = 1\}_{i=1}^{N}$, using the optical algorithm of [11] (which is based on the Hough Transform) for this duality transformation.
>
> Step 2. Filter all intersection points of these lines by thresholding the resulting image at brightness level 2 (these intersection points, and only those points, have brightness at least 2).
>
> Step 3. Dualize the resulting intersection points back into straight lines, using a reverse duality transform, also described in [11]. Clearly, these lines are the desired lines passing through each pair of points in P.

We can extend this technique to the situation where we are given two planar point sets, $P = \{p_i\}_{i=1}^{N}$ and $Q = \{q_j\}_{j=1}^{M}$, and we want to draw on an output SLM all the lines supporting the segments $\overline{p_i q_j}$, for $i = 1, \ldots, N$ and $j = 1, \ldots, M$. In fact, we can also draw just the straight segments connecting between the two given sets (but this requires more optical steps).

Reporting the k Closest Distances among Points in the Plane: Our second example is motivated by pattern matching. If we imagine that we have two images, one being a model, and the other being an image to be matched against the model, and if we further assume that both model and image consist of only finite sets of points (which can be thought of designating special features on the model and image), then a simple pattern matching problem that arises is to find the closest pair of points, or rather the k closest pairs of points among the given points, anticipating that each such close pair designates a match between a model point and an image point. At any rate, this is a basic problem in computational geometry, and we can solve it optically as follows. Let $P = \{p_i\}_{i=1}^{N}$ be the given set of points in the plane.

> Step 1. Construct the Minkowski difference $P - P = \{p_i - p_j\}_{i,j=1}^{N}$.
>
> Step 2. Remove the origin from the image $P - P$, to obtain a new image ΔP.

Step 3. Perform the following chain of conformal coordinate transformations on ΔP:

$$(u,v) \longrightarrow (u^2+v^2, 2uv) \longrightarrow (u^2+v^2, 0) \longrightarrow (\sqrt{u^2+v^2}, 0);,$$

thereby obtaining the set $D = \{(d_{ij}, 0)\}_{i,j=1}^{N}$, where d_{ij} denotes the distance between p_i and p_j.

Step 3. Compute the Minkowski sum

$$S = D + \{(x,0) \mid x \geq 0\}.$$

The optical implementation of the Minkowski sum, as described in [11], has the property that the brightness of each point in the sum is equal to the number of times it is attained by a sum of two points in the source sets. Consequently, as one can easily verify, the brightness in S of the point $(d_{ij}, 0))$ is equal to the rank $rank(d_{ij})$ of d_{ij} in the ordered sequence of distances (each counted with the appropriate multiplicity) among the points of P.

Thus, if we intersect S with D and then filter the resulting image at level k (that is, leave only points with brightness $\leq k$, we obtain the desired k closest distances (as points on the x-axis). We can also find, in a few more steps, the actual pairs of points that attain these distances. Using a symmetric technique, we can also compute the k farthest pairs among the points of P.

References

[1] R. Barakat and J. Reif, Lower bounds on the computational efficiency of optical computing systems, *Applied Optics*, 26(6), 1015-1018 (1987).

[2] R. Barakat and J. Reif, Polynomial convolution algorithm for matrix multiplication with application for optical computing, *Applied Optics*, 26(14), 2707-2711 (1987).

[3] H. Edelsbrunner, *Algorithms in Combinatorial Geometry*, Springer-Verlag, Heidelberg, 1987.

[4] G. Eichman, B. Dong, Coherent optical production of the Hough transform, *Applied Optics*, 22(6), pp. 830-834, 1983.

[5] C. Gu, S. Campbell, J. Hong, Q. He, D. Zhang, P. Yeh, Optical thresholding and maximum operations, *Applied Optics*, 31(26), pp. 5661-5665, 1992.

[6] H. Fujiware, K. Nakagawa, Real-time image substraction and addition using two cross-polarized phase-conjugate waves, *Optics Communication*, 79(1, 2), pp. 6-10, 1990.

[7] J. Reif and A. Tyagi, Efficient parallel algorithms for optical computing with the DFT primitive, *Proc. 10th Symp. on Foundations of Software Technology and Theoretical Computer Science*, 1990, Springer Lecture Notes in Computer Science 472, pp. 149-160.

[8] J. E. Rhodes, Analysis and synthesis of optical images, *American Journal of Physics*, vol. 21, May 1953, pp. 337 343.

[9] M. A. Stuff, J. N. Cederquist, Coordinate transformations realizable with multiple holographic optical elements, *Journal of Optical Society of America*, A 7(6), pp. 977-981 (1990).

[10] A.B. VanderLugt, Signal detection by complex spatial filtering, *IEEE transactions, Information Theory*, IT-10(2), April 1964, pp. 139 145.

[11] Y. B. Karasik, M. Sharir, Optical computational geometry, *Proc. 8th ACM Symposium on Computational Geometry*, ACM Press, pp. 232-241, 1992.

[12] Y. B. Karasik, M. Sharir, The power of geometric duality and Minkowski sums in optical computational geometry, in preparation.

An Algorithm for Implementing Fully-Connected Optical Interconnection Networks with Broadcast Capability

Ahmed Louri
Department of Electrical and Computer Engineering
University of Arizona, Tucson, Arizona 85721
Tel: (602) 621 - 2318, Fax: (602) 621 - 8076

1. Introduction

Broadcasting, or a one-to-all mapping, where a given node (a node in this letter refers to a processor, a memory unit, or a switch) needs to communicate with all other nodes in the system, is a key feature which can support many important activities in parallel and distributed computing. Among these activities are: (1) scheduling activities among different processes and processors in a multiprocessing environment, (2) dynamic load balancing in multiple-instruction multiple-data stream (MIMD) systems, (3) broadcasting instructions from the control unit to all processing elements (PEs) in single-instruction multiple-data (SIMD), (4) maintaining data consistency in caching schemes for multiprocessors, (5) remote procedure calls for distributed processing, (6) access to distributed information, (7) locking and updating in shared-memory systems, (8) parallel numerical computing such as linear algebra algorithms and solution of partial differential equations, and (9) database queries and management. Unfortunately, conventional VLSI technology seems to be unable to implement adequate large-scale interconnection networks with broadcast capability due to the wiring complexity, power, and network latencies. Optics, owing to its inherent parallelism, high spectral and spatial bandwidth, freedom from mutual interference, and minimal timing skew possesses the potential for supporting the communication needs of parallel and distributed computing systems[1].

Several efforts are under way for designing optical interconnection networks for parallel computers. However, many of these proposals do not incorporate the broadcasting function, and the ones that do are either based on the optical matrix-vector[1, 2] multiplication algorithm, or require totally space-variant connection patterns. A possible drawback of the matrix-vector multiplication technique is that it can only be used to connect one-dimensional (1-D) arrays of nodes. In addition, it requires active optical switching to realize the broadcast function. Clearly, optical interconnection topologies designed around two-dimensional (2-D) input and (2-D) output arrays would be more desirable, for they would (1) better utilize the full space-bandwidth product of optical systems, (2) take full advantage of the parallelism of free-space optics, (3) open up new possibilities for designing faster parallel computing algorithms, and (4) would be better suited for the recent advances in compact, two-dimensional optical and opto-electronic logic devices (OEICs). While totally space-variant interconnection architectures can provide arbitrary connection patterns and possibly broadcasting, they not only put stringent demands on space-bandwidth product (SBWP) but also require complex optical hardware for their implementations.

2. Proposed Algorithm

In this paper, we present an algorithm for constructing 2-D to 2-D free-space optical broadcast interconnection networks that does not require totally space-variant optics. The algorithm performs broadcast function in constant time regardless of the size of the input/output arrays. Fig.1 illustrates the algorithm. For the sake of clarity and without any loss of generality, we use input/output arrays of 4×4 nodes. Assume that each node has associated with it an optical source for signal transmission and an optical detector for signal detection. These sources are assumed to be mutually incoherent. Given an input array as shown in Fig.1.a, the first step of the algorithm consists of vertically deflecting and superposing all the rows of the array onto a given row as indicated in Fig.1.b. Any row of the array can be used for superposing the rest of the rows on it. The resulting image is then replicated four times to fill the size of the original array. In general, the number of replicas is equal to the number of rows comprising the source array. The result of step two is shown in Fig.1.c. Note that after step two, any node can communicate with all the other nodes residing in the corresponding column. Next, the columns of Fig.1.c. are deflected horizontally and superposed to form a single column as shown in Fig.1.d. Finally, the resulting column is replicated four times to fill the size of the original input array as shown in Fig.1.e. In the final image, every node communicates with every other node in the input array as indicated by the notation $1 \ldots 16$ in each node position. In fact the four-step algorithm summarized in Fig.2 (vertical deflection and superposition, replication, horizontal deflection and superposition, and replication) achieves full connectivity between nodes, and as a result of this, it also implements the broadcasting function. Optical implementations of the proposed algorithm require relatively simple optical hardware. The two basic operations required are row/column deflection and superposition, and row/column replication. These operations can be implemented using a wide variety of optical components including classical optics hardware such as prisms, lenses, beam splitters, mirrors, etc., grating elements such as blazed gratings, and holograms[3]. One possible optical implementation would be the use of two holographic optical elements (HOEs) and two Fourier lenses arranged in a 4f system. For an input array of $n \times n = N$ nodes, the first HOE would be composed of $n = \sqrt{N}$ subholograms, each of which is adjacent to each row of the input array. Each subhologram would have a linear grating with a period and orientation that is different from the rest of the subholograms. The first HOE, along with the first Fourier lens directs all the nodes within each column to the same location in the Fourier plane. This arrangement implements the first two steps of the algorithm. The second HOE would be located in the Fourier plane of the first lens, and is also composed of $\sqrt{N}$ subholograms. This HOE along with a second Fourier lens would implement the last two steps of the algorithm.

We now estimate the space-bandwidth product of proposed algorithm without spatial addressing. According to Garbor's theorem[4], the number of degrees of freedom (F) in a linear optical system can be expressed as

$$F = A\Omega/\lambda^2 \tag{1}$$

where A is the area of the optical beam, Ω the solid angle of the beam spread, and λ is the wavelength of light. For an optical interconnect system with N pixels and M different connection patterns (or point-spread functions), the number of degrees of freedom can also be expressed as:

$$F = M \times N \tag{2}$$

Suppose that the input array is a square with ($n \times n = N$) nodes. As can be seen in Fig. 2, the proposed algorithm requires n connection patterns at step one to superpose n rows into a single row, one connection pattern at step two to replicate the superposed row, n patterns at step three to superpose n columns into a single column, and one pattern at step four to replicate the superposed column. Thus, the algorithm requires no more than $n = \sqrt{N}$ different types of connection patterns. Therefore, from Eqs. 1 and 2, the maximal SBWP capability of the proposed algorithm is:

$$N = (A\Omega/\lambda^2)^{2/3} \tag{3}$$

We compare this result with a fully-connected network based on space-variant interconnects. The number of different types of connection patterns is N for a space-variant fully-connected netwwork, which result in maximal SBWP of:

$$N = (A\Omega/\lambda^2)^{1/2} \tag{4}$$

From Eqs. 3 and 4, we see that the proposed algorithm can potentially offer a substantial improvement in SBWP. For example, with an input array of size $50mm \times 50mm$ and wavelength 0.5 μm, the maximum number of nodes obtained by the proposed algorithm is 4,600,000, while the space-variant method can only support 100,000 nodes.

In summary, we presented an algorithm that performs full connectivity between nodes arranged in planer fashion. The fully connected nature of the algorithm allows for broadcasting, where one node can send information to the rest of the nodes in the array. This very desirable feature is absent or limited in many optical interconnection networks proposed so far. The algorithm can offer a substantial SBWP improvement as compared with space-variant interconnects. The implementation requires two basic operations (image superposition and replication) which are highly amenable to optical implementations with relatively simple optical hardware.

References

[1] A. A. Sawchuk, C. S. Raghavandra, B. K. Jenkins, and A. Varma, "Optical crossbar networks," *IEEE Computer*, vol. 20, no. 6, pp. 50 - 62, June 1987.

[2] A. R. Dias, R. F. Kalman, J. W. Goodman, and A. A. Sawchuk, "Fiber-optic crossbar switch with broadcast capability," *Optical Engineering*, vol. 27, no. 11, pp. 955 - 960, Nov. 1988.

[3] B. K. Jenkins, A. A. Sawchuk, T. C. Strand, and B. H. Soffer, "Sequential optical logic implementation," *Applied Optics*, vol. 23, no. 19, pp. 3473 - 3474, Oct. 1984.

[4] D. Gabor, "Light and information," *Progress in Optics*, vol. 1, pp. 109 - 153, 1961.

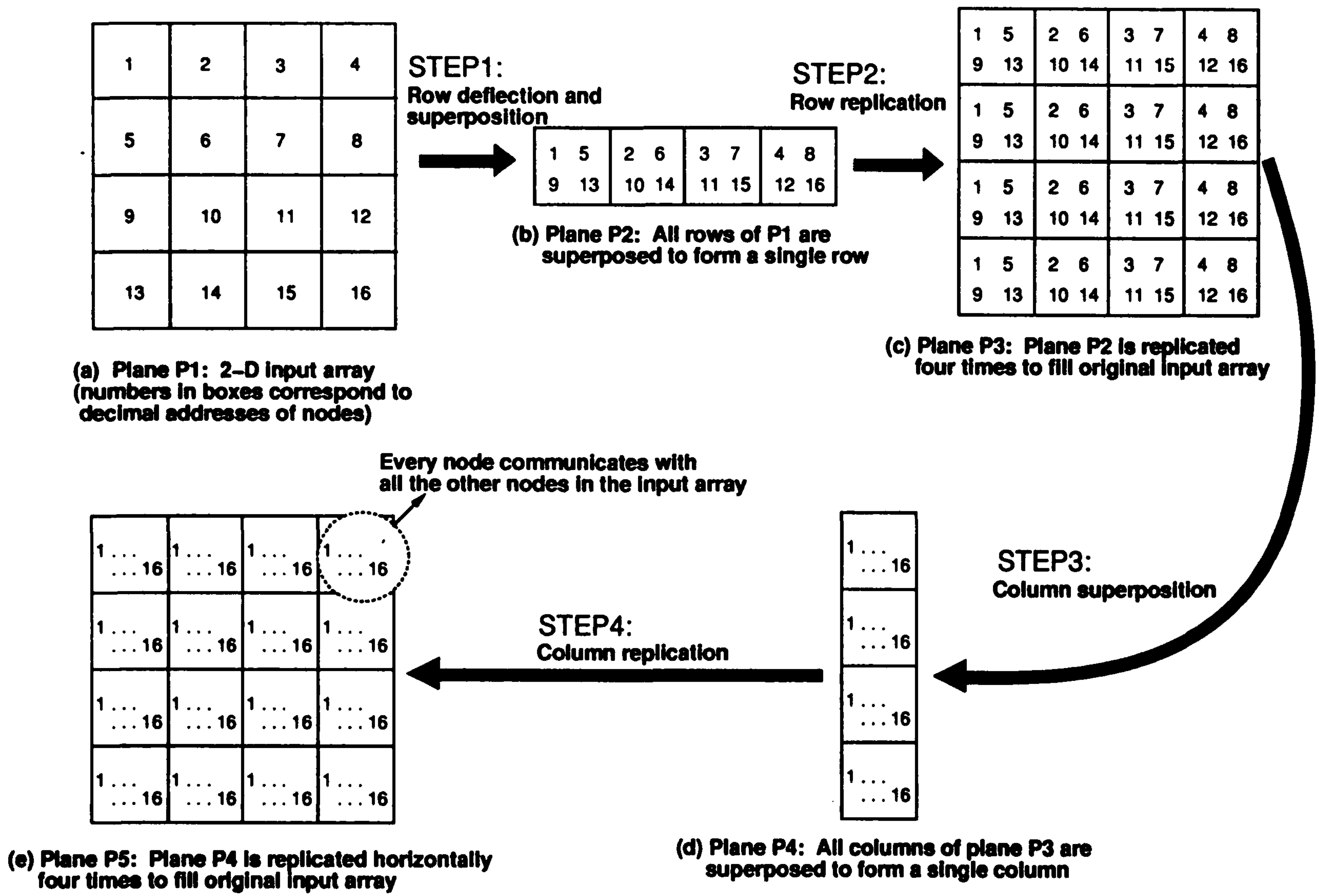


Fig 1: An algorithm for designing fully–connected optical interconnection networks with broadcast capability

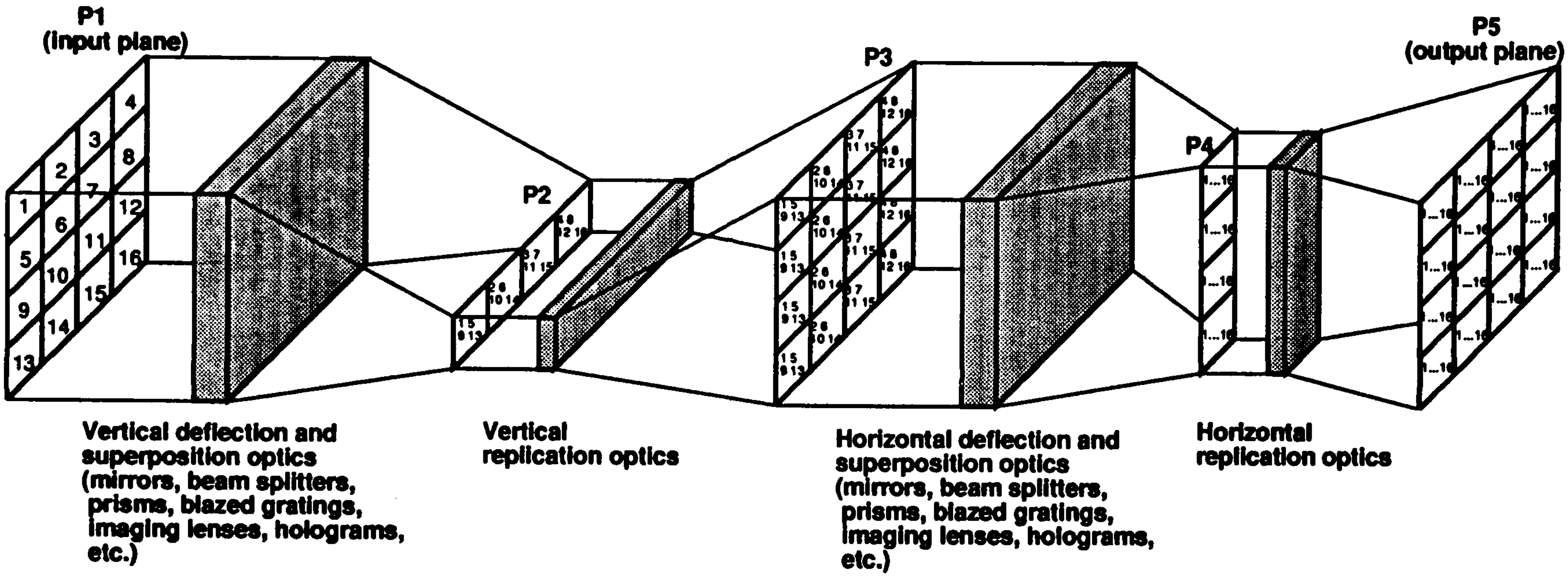


Fig 2: An optical architecture for the algorithm

An Optical Data Filter for Selection and Projection

P. A. Mitkas, S. A. Feld, L. J. Irakliotis, and C. W. Wilmsen
Center for Optoelectronic Computing Systems
and Department of Electrical Engineering
Colorado State University
Fort Collins, CO 80523

1. Introduction

The management of today's very large databases requires high storage capacity and parallel searching to ensure real–time operation. The two main characteristics of database processing are: a) most of the transactions are based on comparisons of the data fields with a search argument, and b) a large amount of data have to be retrieved and processed in order to generate the query result, which is usually a small fraction of the database. Therefore, a common solution to the problem is a special purpose processor operating as a data filter. This unit receives records from the database and screens them on the fly according to user defined criteria. If a record satisfies the query, it is transferred to the end user; otherwise, it is rejected.

Recent advances in optical technology have yielded optical memories, such as parallel read out optical disks and holograms, that offer both large storage capacity and massive transfer rates, two essential ingredients in every database environment [1]. In contrast to the serial output of most electronic and magnetic media, the output of parallel optical memories can be one– or two– dimensional and reach hundreds of megabytes per second. Although desirable, these data rates can overwhelm electronic computers which are designed to operate at a few megabytes per second. Fortunately, the nature of the above applications can provide a solution because most of the operations involve searching through a large data volume to identify and retrieve data that satisfy certain selection criteria and match a reference pattern. This unique behavior suggests that the effective data rate from optical storage to the electronic host can be significantly reduced if a parallel optical processing unit is employed to perform some first order filtering operations. The availability of fast and highly parallel optical processing elements [2] makes the application of optical techniques to database processing a viable alternative [3]. In this paper, we present the architecture and functionality of an optical database filter based on arrays of LAOS optical gates. The filter is performing selection and projection operations in a relational database environment.

2. Selection and Projection in a Relational Database

Data in a relational database are stored in two–dimensional arrays called relations. A database may comprise multiple relations. Each row in the array corresponds to a record or tuple, while the columns demarcate the various data domains or attributes. Every relation has one or more domains that serve as unique identifiers for the tuples and are called primary keys. A set of relational operations are defined and can be used to extract information from relations. Most of them select rows and/or columns of a data array according to certain criteria. There are two main categories: a) the traditional set operations (a relation is a set whose elements are tuples) which include union, intersection, set difference and Cartesian product, and b) the relation–oriented operations, such as projection, selection and join. Projection operation enables a user to select columns from a relation and specify their order. Selection is the process of retrieving tuples from a relation that satisfy one or more selection criteria. The data entries in certain fields of a record must be compared against the user supplied selection arguments. If a match is detected, the record must be retrieved.

3. The Light Amplifying Optical Switch (LAOS)

The architecture of the optical data filter is based on a photonic switching device called the Light Amplifying Optical Switch (LAOS). The LAOS is a phototransistor vertically integrated with a light–emitting diode (LED). Its upper half is a heterojunction bipolar transistor (HPT) that acts as an optical detector and provides gain. The lower half of the device is a double–heterojunction light– emitting diode that is driven by the phototransistor and provides an optical output signal [4]. The vertical structure of the LAOS relaxes the need for ultra–fine–line lithography, and allows the LAOS to be densely packed into large 2–D arrays. Typical electronic logic circuits consist of multiple input stages that invert and restore the input signals. The LAOS can be used to implement optical logic circuits with similar characteristics. Several optical gates based on the LAOS have been demonstrated, including inverters, AND, OR, NOR, XOR, latches and flip–flops [5]. Two–dimensional arrays of NOR gates have been fabricated and are currently being evaluated. Furthermore, we have successfully fabricated microlenses at the output of these optical gate arrays making them cascadable.

3. Selection Operation on a 2–D Data Array

The selection operation can be described as choosing those tuples of a relation that have a field or fields which match some given search argument. The operation can be broken down into two steps: a) masking–off all of the unwanted fields, and b) performing a bitwise XOR between the search argument and the remaining fields of the tuples. The first step can be performed on the input array using a 2–D array of optical AND gates, and the second step by cascading the output into an array of optical XOR gates, as shown in Figure 1.

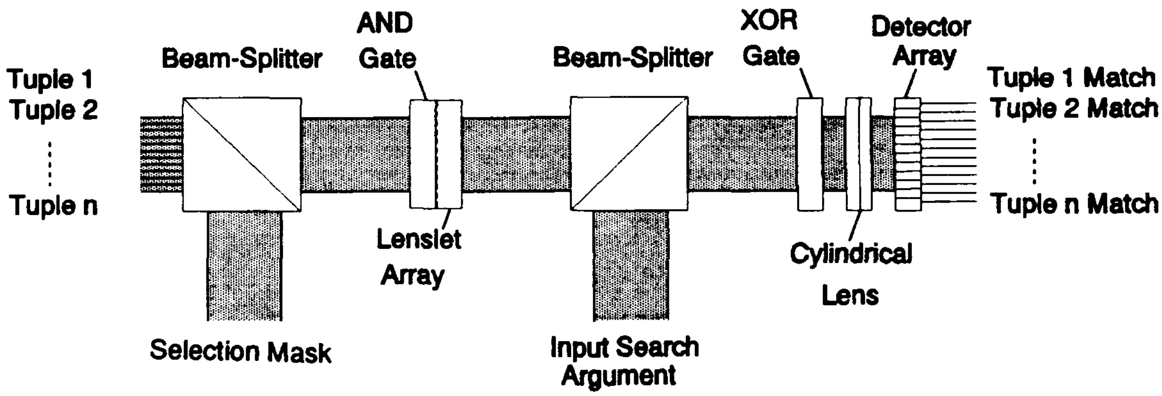


Figure 1. The selection module.

Here, the outputs of the XOR gates for each tuple in the array are ORed together by the cylindrical lens and imaged onto the linear photodetector array. The outputs of the photodetector array are used to enable the tuples that satisfy the selection criteria. The selection mask contains 1s in every bit position of the desired fields and 0s for the rest. This pattern is replicated in n rows of an SLM and brought to the selection module through the first beamsplitter. The data array contains one tuple per row and is interleaved with the selection mask to form the input to the 2–D array of AND gates, the pixels of which are shown in Figure 2.a. An H–cell in this diagram corresponds to an HPT, while each L–cell is a LAOS device. Note that the only electrical connections needed for each pixel are the power rails, with all data signals being transferred normal to the array.

The output of the AND gate array is propagated through the transparent substrate to an array of micro–lenses. The focal length of this array of lenses is designed such that they will focus the output spots of the AND gate onto the next array of logic gates. Using a second beamsplitter, the outputs of the AND gate array are mixed with the array of selection arguments coming from another SLM. The output of the beamsplitter is then directed onto the array of XOR gates, the pixels of which are shown in Figure 2.b. The outputs of the XOR pixels are brought into a cylindrical lens which is oriented such that all of the outputs of a single row are incident upon a single detector of a linear detector array. The outputs of the photodetectors are used to determine which tuples match the search argument, and are used in the projection operation, which will be discussed in the next section. Typically, a large number of data pages will be processed before a new set of selection masks and search arguments must be loaded into the SLMs.

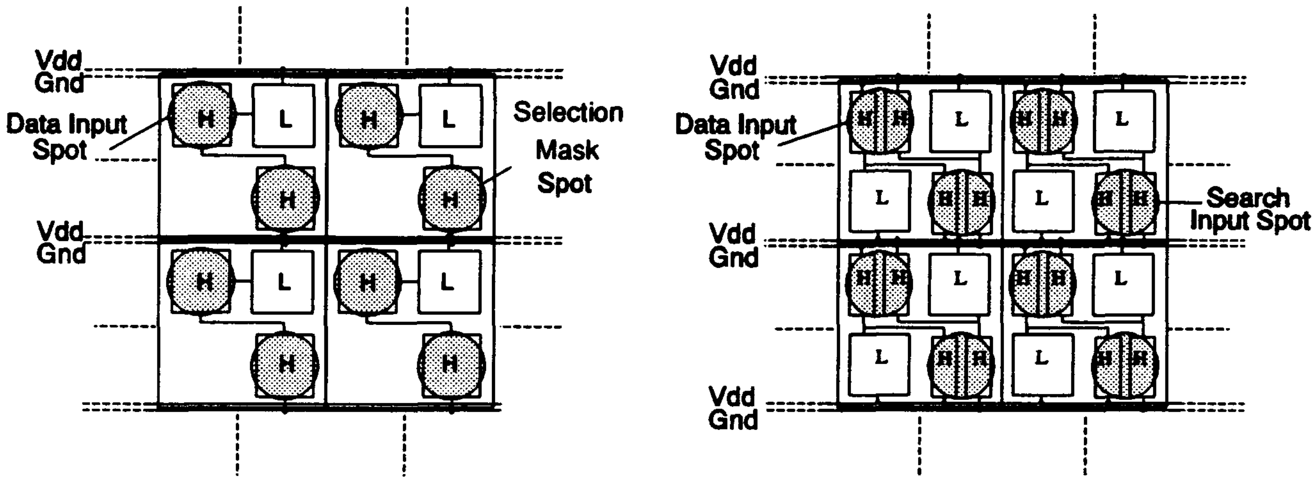


Figure 2. Pixel diagrams of a) the optical AND gate array and b) the optical XOR gate array. In each pixel, the two shaded areas show the positions of the two input beams.

5. Projection Operation on a 2–D Data Array

The projection operation was essentially performed during the first half of the selection operation described in the previous section. However, during the selection operation, the results of the projection were processed in parallel by the XOR gate array. If the output of the database filter were to go to an optical parallel processor, then the results of the projection could be left in a 2–D optical format. However, for our case, the output of the projection operation must be converted to an electronic signal and stored in electronic memory for subsequent transfer to an electronic processor. Recently, there has been work done to combine arrays of optical detectors monolithically with large memory arrays [6]. The optical memory described in [SAY91] could be written in parallel in 0.78 ns, and read in times comparable to those of standard SRAMs. The outputs of the selection module can be used to select the proper tuples from the optoelectronic memory which would subsequently be transferred to the processor, as illustrated in Figure 3.

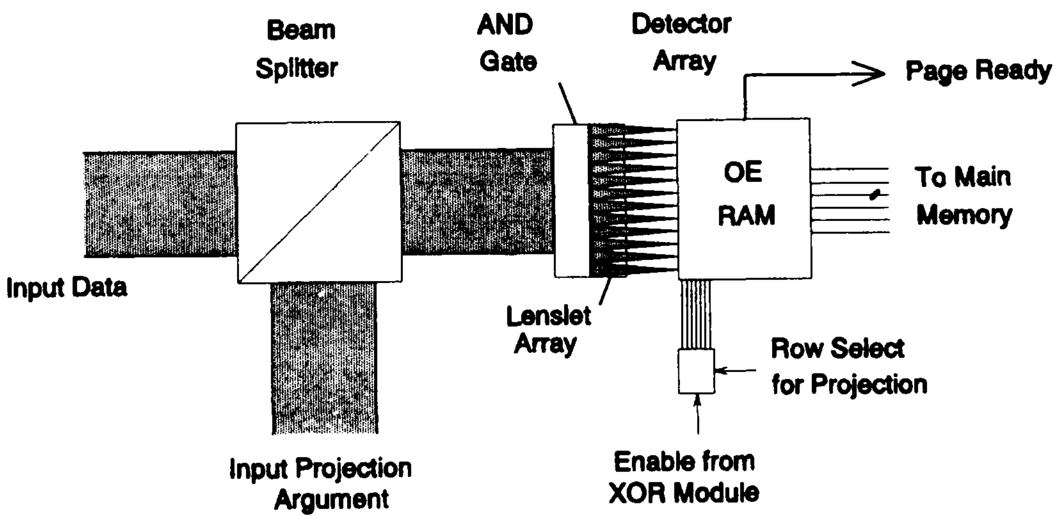


Figure 3. The projection module.

6. The Optical Data Filter

The block diagram of the optical data filter is shown on Figure 4. Data are input as 2–D pages from the left. The bottom channel is the selection module while the upper channel performs projections. The steps involved in a projection/selection transaction are illustrated with an example.

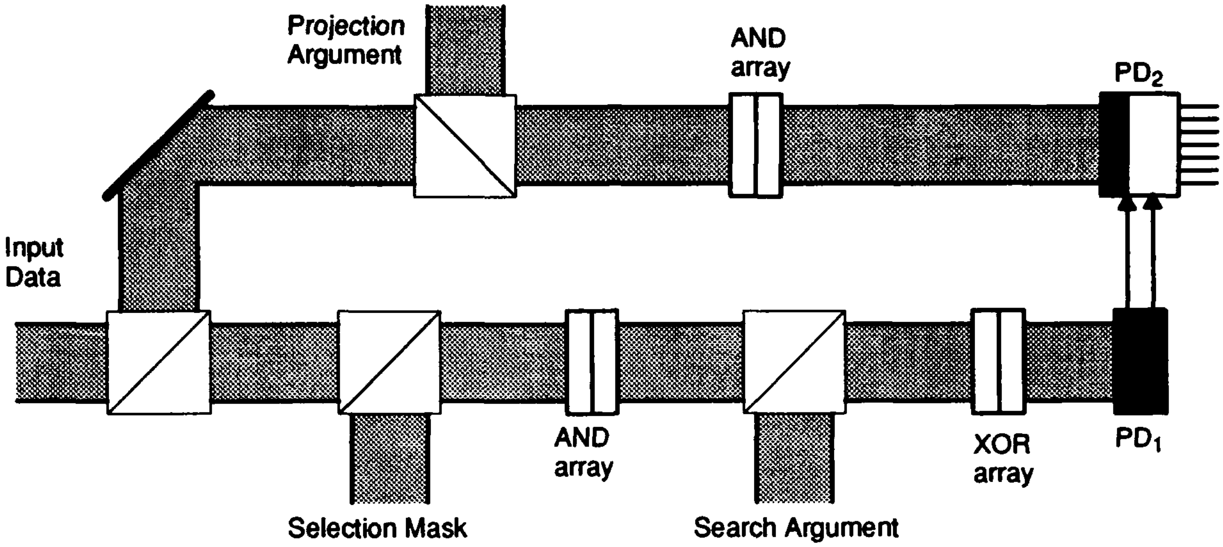


Figure 4. Block diagram of the optical data filter.

Suppose that a relation in the database of a bank that issues a credit card contains four data fields: *Transaction_Id* (30), *Customer#* (25), *Purchase_Date* (15), and *Price* (20), where the numbers in parentheses indicate the length in bits of each field. Then a request to retrieve the numbers of the customers who made a transaction on November 5th, 1992 will proceed as follows:

a) The selection mask will contain 1s in bit positions 55–69, because the selection is based only on the contents of the *Purchase_Date* field. The remaining bit positions will have 0s.

b) The encoded value of 11/5/92 will be loaded in bit positions 55–69 of the search argument with the rest of the bits being zeros.

c) The projection mask will contain 1s in bit positions 30–64 because we are interested in recovering only the *Customer#* fields.

After all the query arguments are loaded, data pages are read in from an optical memory. Rows are selected by the output of PD_1 and columns are removed at the output of the AND array on the upper channel. The filter can be programmed to perform a projection without selection and vice versa.

The response time of the filter is proportional to the switching times of the optical gates, which are currently 60 ns and are expected to reach 10 ns.

7. References

1. P.A. Mitkas and P.B. Berra, "Parallel Optical Memories for Very Large Databases", Proc. of 37th SPIE Annual Symposium, June 1992, SPIE Vol. 1773C.
2. C.W. Wilmsen, et al, "Switching Light with Light", IEEE Circuits and Devices, Vol. 7, No. 6, 1991, pg. 21.
3. P.B. Berra, A. Ghafoor, P.A. Mitkas, S.J. Marcinkowski, M. Guizani, "The Impact of Optics on Data and Knowledge Base Systems", IEEE Trans. Knowledge and Data Eng., Vol. 1, No. 1, March 1989, pg. 111.
4. S.A. Feld, et al., "Electrical and Optical Feedback in an InGaAs/InP Light–Amplifying Optical Switch (LAOS)", IEEE Trans. Elect. Dev., Vol. 38, No. 11, Nov. 1991, pg. 2452.
5. C.W. Wilmsen, F.R. Beyette, Jr., X. An, S.A. Feld, K.M. Geib, "Smart Pixels using the Light Amplifying Optical Switch (LAOS)", to be published, IEEE J. Quant. Elect. Special Issue on Smart Pixels, Feb. 1993.
6. A.H. Sayles, J.P. Uyemura, "An Optoelectronic CMOS Memory Circuit for Parallel Detection and Storage of Optical Data", IEEE J. Solid State Circuits, Vol. 26, No. 8, Aug. 1991, pg. 1110.

Polarization and birefringence based optical system for reconfigurable weighed interconnects and image shifting

Werner Peiffer, Hugo Thienpont, Michiel Pelt,
Roger Vounckx and Irina Veretennicoff.
Vrije Universiteit Brussel, Applied Physics Department TONA-TW
Pleinlaan 2, B-1050 Brussel, Belgium
tel. 32 2 641 3567 / fax 32 2 641 3450
telex VUBCO-B 61 051

Three key operations are sufficient in order to construct a cellular automaton: *routing* and *shifting* of information contained in 2-D data planes, and the *logic NOR* operation applied on the data [1].
In this paper we focus on the shifting operation for which we recently proposed a processor subsystem that provides the facilities to perform the required dynamically reconfigurable nearest neighbour interconnects [2]. We call this element a *Dynamic Shifter* since it enables to interconnect in a dynamic way all pixels from one logic plane to one of their nearest neighbours on the next logic plane. This is done by shifting the information contained in the first data plane in one of the eight possible nearest neighbour directions.
The principle of operation of this shifter is based on a combination of *Birefringent Crystals, Large Aperture liquid crystal Retarders* (LAR)and *polarised light beams.*The dynamic shifter consists of four modules, each of which is made up of a LAR and a birefringent crystal. The modules are oriented and stacked in such a way that a shift can be performed in any of the eight possible nearest neighbour directions, depending on the linear state of polarization of the input beams (TE or TM) and on the PC-controlled voltages of the four LARs (figure 1).
In the first part of this paper we will discuss and demonstrate in detail several possible applications of this system in the field of 1-to-1 and 1-to-many interconnects. 1-to-1 interconnects are most suited when the optical logic gates used in the optical processor have a low fan-out. In that case, using the dynamic shifter, all the pixels from one logic plane are interconnected to only one of their nearest neighbours on the next logic plane (figure 2). In this way a large number of space-invariant parallel information channels can be realised and configured dynamically, enabling the execution of non-directional as well as directional cellular algorithms. The degree of parallelism is completely determined by the dimensions and the pitch of the data planes addressed by the shifter. The system itself, i.e. the combination of the LARs and the birefringent crystals, is not imposing an upper limit to the resolution due to the fact that the LARs are not pixelled. This can be clearly seen in the case of high resolution image shifting, where the degree of parallelism corresponds to the photographic resolution of the image being processed. At the conference site we will illustrate our latest results in this field (videotaped).
In the domain of 1-to-many interconnects we demonstrate the possibility of weighed interconnections. We here make distinction between symmetric and asymmetric weight distributions which may be useful e.g. for the implementation of optical neural nets. In figure 3 we have depicted several weight distributions with their experimental verification in the case of a single input beam. Applying complete images to the shifter -configured for weighed interconnects- will result in filter operations. For example a binomial filter, which has an image smoothing effect [3], can be implemented with the weights depicted in figure 4a. Figures 4b-d show the effect of this filter on a single input beam, on a draught-board pattern and on a picture of a sharp edge. The real-time nature of the filtering operations has been videotaped and will be shown at the conference.
In the second part of the paper we treat the combination of a diffractive optical 8x8 beamgenerator with the shifter system. We show, making use of quantitative measurements of the diffraction efficiency and the beam profiles, that this approach makes it possible to address our silicon-based optical logic planes [4, 5].

References

1. W. Peiffer and H. Thienpont, 'An optical 2-D binary cellular automaton for all purpose optical computing', Proceedings of Les 7èmes journées d'étude *Fonctions optiques dans l'ordinateur*, Brest, (1991).

2. H. Thienpont, W. Peiffer, M. Pelt and I. Veretennicoff, "Optical module for dynamically reconfigurable nearest neighbour interconnects", in *Proceedings of the ICO Topical Meeting on Optical Computing, SPIE Vol. 1806*, Minsk, Republic of Belarus.

3. T.J. Fountain and V. Goetcherian, 'CLIP 4 Parallel Processing System', *IEE Proc.* **127**, Pt. E, 5, pp. 219-224 (1980).

4. H. Thienpont, I. Van De Voorde, I. Verbesselt, W. Ranson, C. De Tandt, R. Vounckx and I. Veretennicoff, "Hybrid electrically assisted thermo-optic resonators for prototype digital optics: from element to array", *Optical Computing and Processing*, 1 (2), pp. 145-155 (1991).

5. H.Thienpont, W. Peiffer, W. Ranson, C. De Tandt, R. Vounckx, I. Veretennicoff, A. Koster and S. Laval, 'Silicon based optical logic planes: towards a second HEATOR generation', submitted to this conference.

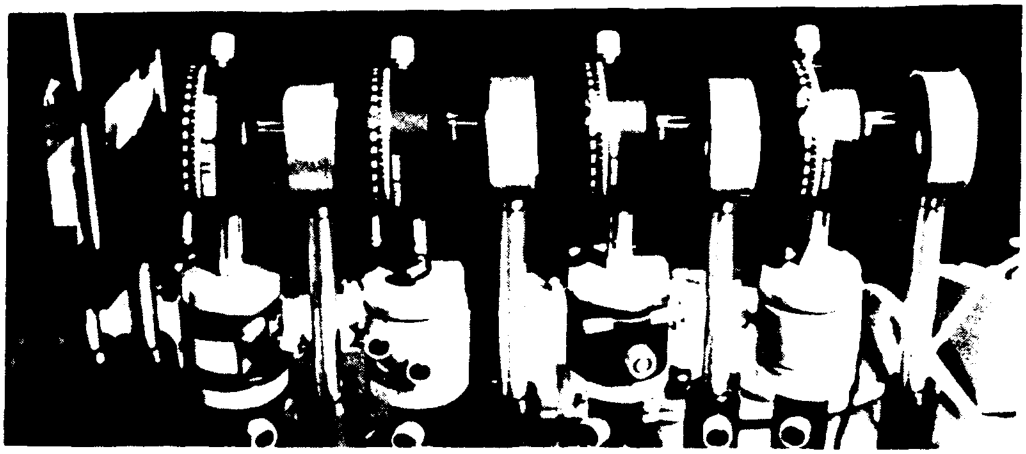


Figure 1: Picture of the dynamical shifter setup.

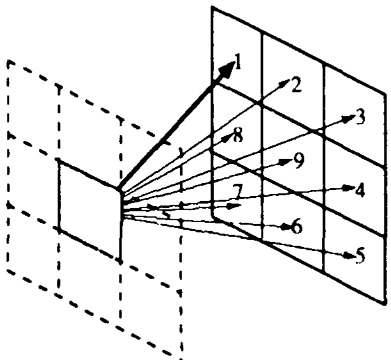


Figure 2: The dynamical shifter enables dynamical reconfigurable interconnects in the eight nearest neighbour directions. The interconnection in direction 1 is depicted in bold.

0	**0**	**3/4**
0	**0**	**1/4**
0	**0**	**0**

0	**1/4**	**0**
0	**1/4**	**1/4**
0	**0**	**1/4**

0	**0**	**1/3**
0	**0**	**1/3**
0	**0**	**1/3**

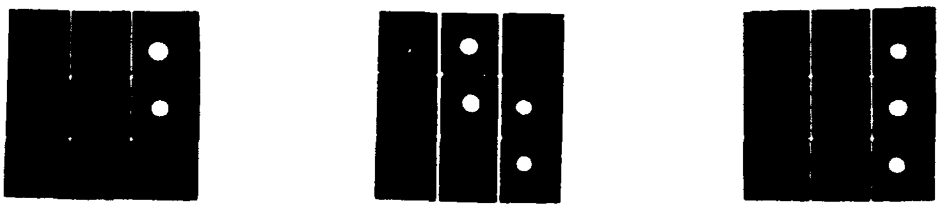


Figure 3: Weight distributions and experimental verification in the case of a single ingoing beam.

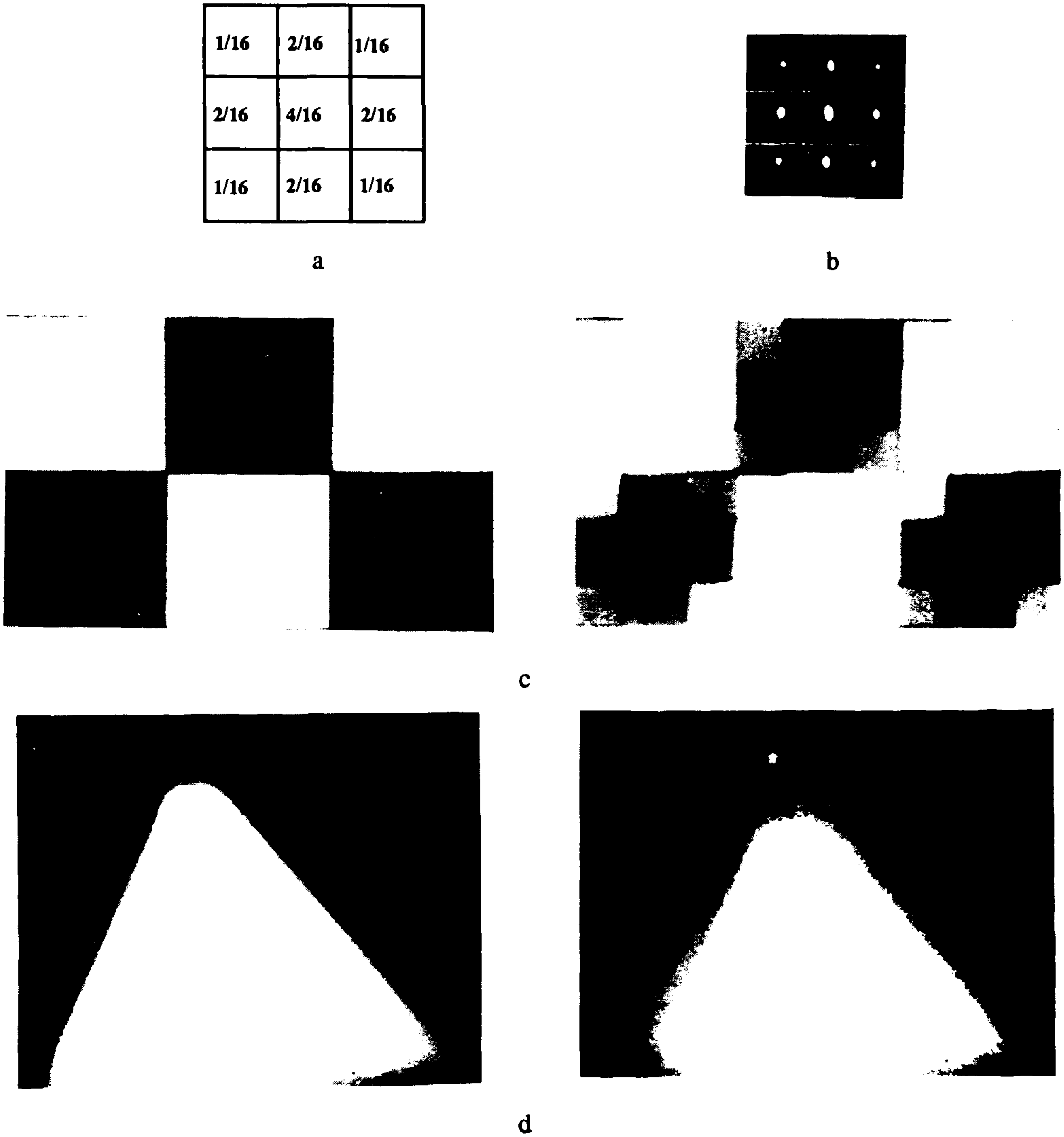


Figure 4: a. Binomial weight distribution;
b. Binomial filtering of a single beam;
c. Original and binomial filtering of a draught-board pattern;
d. Original and binomial filtering of a sharp edge.

Optical interconnection slab based static dataflow computer

Alastair D. McAulay

Lehigh University
Department of Electrical Engineering and Computer Science
19 Memorial Drive West
Bethlehem, PA 18015, (215) 758-4070

Future high speed computers must be physically small and use optical interconnections in order that data movement keep up with increasing electronic processing element speeds. Japanese MITI has a 10 year goal of demonstrating a computer capable of 1 teraflop/sec/cm^3. This paper proposes an architecture and design that could lead to such performance. Static dataflow is selected because it is simpler to construct and has wide applicability where the same numerical algorithm is run frequently or even continuously. The benefits of dataflow are reviewed in references [1] [2]. The proposed architecture uses optical interfacing to processor chips to overcome limits with electronic interfacing of approximately 10,000 pins and 1 Gb/s per pin [3]. It uses flip chip bonding of electronic VLSI chips and GaAs laser diodes onto a transparent substrate for optical switching. This was previously demonstrated [4] [5] and discussed at the 1992 Annual OSA Meeting which is reviewed in reference [6].

Description of proposed optical dataflow computer

The proposed architecture, shown in figure 1, is based on an earlier design by the author [7] [8]. At that time the design was uneconomic because every processor needed a laser diode costing approximately $600 each and optical busses were expensive. Today, microlasers [9] are relatively inexpensive so that the design is practical. Similarly, there was no viable optical switch at that time. Today, slabs of transparent material have been shown to be effective [4] [5]. The design will accommodate a hundred electronic processors connected into an optical interconnection network. Only nine processors P_1 through P_9 are shown for simplicity of explanation. The outputs from the switch are connected back to the inputs of the processors. Processors also have outputs and inputs to a memory unit which is constructed similarly.

An algorithm is mapped to a directed graph using one of the language and compiler developments for dataflow machines. The level of parallelism may be observed from the flow graph and may be increased by manipulating the algorithm [10]. Computation and control nodes in the flow graph are assigned to processing elements in the system and links in the flow graph to settings of the switch. Data flows into the switch during operation and is routed to the appropriate processor. A processor will perform the operation for which it is programmed on the next clock cycle after receiving its operands. The processors have local memory including stacks, queues, buffers, internal busses, and local random access memory. Parallel to serial converters drive microlasers on the processor chip to connect to the optical switch. The output is routed via the switch to the next processor. Dataflow architectures tend to reduce overhead associated with instruction decodes, address computation, and data fetch and store.

Figure 2 shows the design for the processing unit having a 3×3 array of processors for illustration. The use of GaAs chips allows microlasers to be constructed directly on the chip. If silicon VLSI processor chips are used, the GaAs laser diodes can be flip chip mounted alongside the silicon chips. Each chip has a microlaser and two light detectors mounted in the x-direction and another set in the y-direction. The microlasers and detectors are arranged in different positions for each chip in the x-direction, providing a unique channel for broadcasting from each chip. The beams emitted by the microlasers are shown darker. The beam expands in the glass and strikes a binary optic computer generated hologram etched into the lower surface of the glass using binary optics. The light is collimated at an angle. Researchers have demonstrated 89% diffraction efficiency at 27^0 [4]. The beam passes through air and strikes a mirror. On reflection, it is diffracted by another hologram etched into the glass that causes part of the light to be focussed into a neighbor chip. Only a small amount of light is needed because it is focussed on the detector. The remainder of the light continues as if reflected from a mirror in order to reach subsequent processors. The same procedure is followed in the y-direction so that a processor can communicate with any other in less than two steps or in less than 300 ps for a 10×10 processor.

A set of fiber optic ribbon cables along one edge, as shown, allows light to be propagated off the unit to the memory unit constructed similarly but allowing optical interconnection only in the y-direction. An electronic conditioning interface chip may be used with the optical ribbon fiber to ensure quality signals.

Simulation results

A simulator was developed to demonstrate the concept, and to permit testing of parallelism for different algorithms. We consider an algorithm with iterative loops, in contrast to previous research in which iterative loops were unrolled [8]. Figure 3 illustrates a dataflow graph for the computation $z = x^n$ [11]. The right hand loop counts to n and provides the control for the other two loops. The left hand loop circulates x. The center loop allows a *one* to pass through merge unit M_2 at the start when the control line is *false*. It then circulates the value multiplied by x in F_1 at each iteration. When the control line switches back to *false* at the end, the result is passed out through G_4.

Computer code representing the flowgraph was written to provide input to the simulator [8]. The simulator generates an activity chart, figure 4, indicating active nodes at each time step of the clock for a three pipeline x^2 computation followed by four iterations of a three pipeline x^n computation.

Conclusion

Simulation, experiments, and analysis suggest that the proposed relatively easy to construct machine should compete favorably with all electronic processors and provide potential for much faster machines than are possible with electronics alone.

References

1. J. Dennis, "The evolution of static dataflow architectures," Chapter 2 in *Advanced topics in dataflow computing*, Eds. J. Gaudiot and L. Bic, Prentice Hall, 1991.
2. A. L. Davis and R. M. Keller, "Dataflow programing graphs," IEEE Computer, **15**(12), pp 26-41, 1982.
3. H. S. Hinton, "Smart pixels and free-space interconnection," Byte, **17** (9), p 174, Sept. 1992.
4. M. Feldman, Univ. of North Carolina at Charlotte, Personal communication, 1992.
5. J. Jahns, AT&T, Personal communication, 1992.
6. A. D. McAulay, "Innovations in optical computing,", Computer in Physics, Nov/Dec. 1992.
7. A.D. McAulay, "Optical Crossbar Interconnected Signal Processor with Basic Algorithms," Optical Engineering, **25**(1):82-90, 1986.
8. A. D. McAulay, *Optical computer architectures*, (John Wiley, New York, 1991).
9. J. L. Jewel, J. B. Harbison, and A. Scherer, "Microlasers", Scientific American, **265** (5), 86-94, 1991.
10. A.D. McAulay, "Conjugate Gradients on Optical Crossbar Interconnected Multiprocessor," Journal of Parallel and Distributed Processing, **6**:136-150, Feb. 1989.
11. K. Hwang and F. A. Briggs, "Computer architecture and parallel processing," McGraw-Hill, 1984.

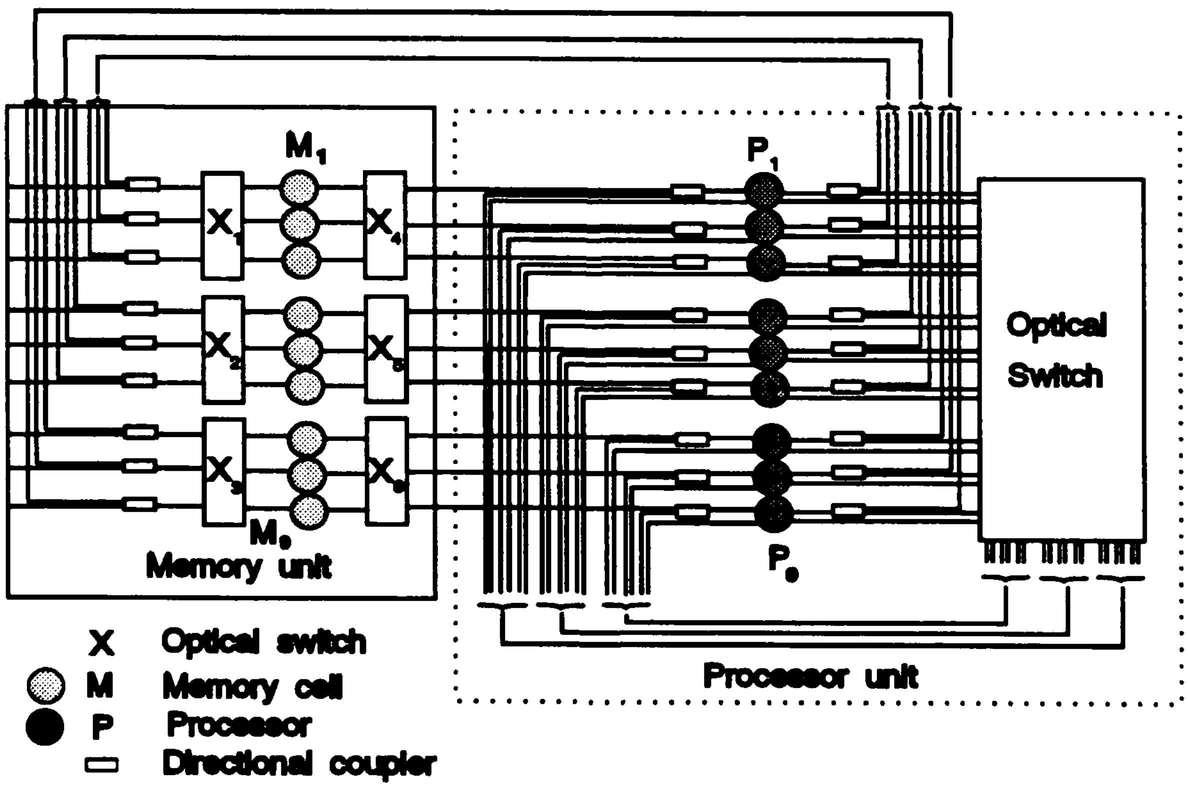


Figure 1: Block diagram of optical dataflow computer

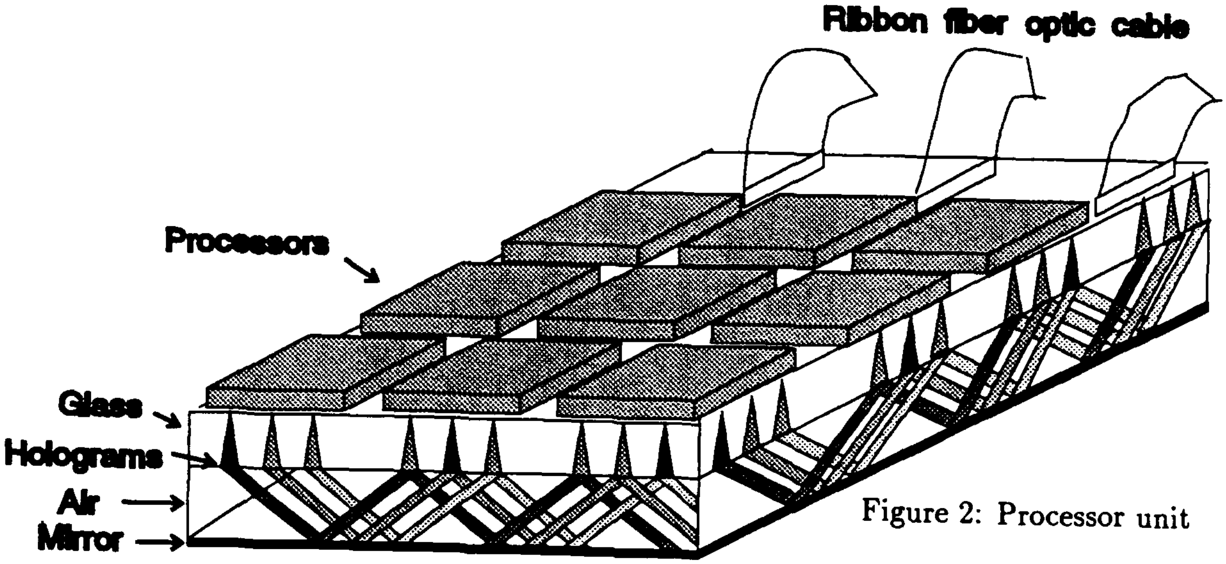


Figure 2: Processor unit

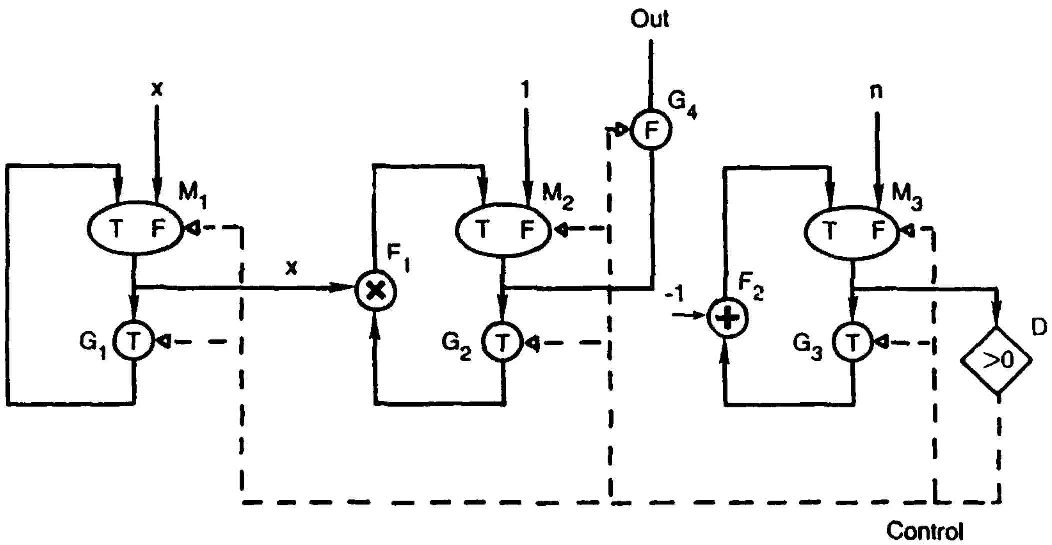


Figure 3: Dataflow graph for $z = x^n$

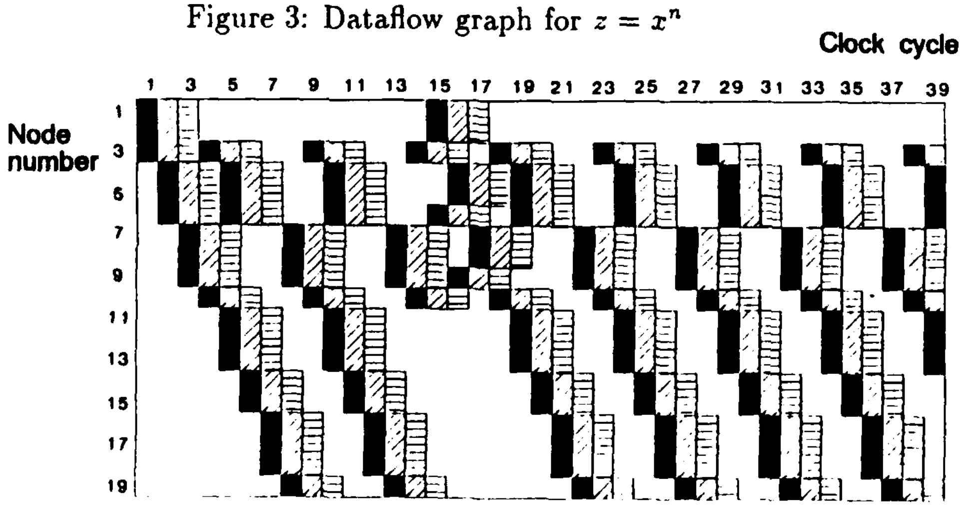


Figure 4: Activity Chart for dataflow graph for figure 3

Board-to-board high-speed optical interconnections

Jie Jiang
Department of Communication Engineering, Duisburg University
Bismarckstr. 90, D-4100 Duisburg 1, Germany
Phone: +49-203-379-2547

1. Introduction

Telecommunication systems using optical fibers for high speed data transmission have been well accepted for long term. The deployment of cptics in local area networks (LANs) for machine-to-machine interconnections is also more and more standard. Thinking consequently, optics should be applied next to board-to-board interconnections.

But optical board-to-board interconnections may be very different from the well known optical fiber technologies. The optical or optoelectronic components like fibers, connectors, star couplers, discrete optical sending and receiving diodes and their driving and amplification circuits, respectively, are unsuitable for board-to-board interconnections. The data transmission on the board-to-board level often is bi-directional and broadcasting (point to multi-points).

Nevertheless optics generally offers many advantages in interconnection technology, especially on the board-to-board level [1], also called backplane level.

In this paper a new concept using a cylinder mirror and especially formed light-guiding plates is introduced. This concept is appropriate for board-to-board high-speed interconnections in multi-processor systems. Because of the similarity to another concept introduced by the author [2] this optoelectronic backplane is called Optical Parallel Plate Stack with Cylinder Mirror (OPPSCM). Some measurement and simulation results concerning the optical power transmission are presented, too.

2. Concept of the optoelectronic backplane OPPSCM

In a computer system an electronic backplane is used generally for data transmission between several boards (Fig. 1). In order to profit by optics without many changes on the electronic boards, laserdiodes (LDs) with their driving circuits and photodiodes (PDs) with their amplification circuits are applied as optical data inputs and outputs, respectively, mounted at the edges of the boards. They replace the electronic transceiver components and the electronic connectors in an electronic backplane system. Ideally, LDs and PDs are integrated with their driving and amplification circuits, forming 1-dimensional arrays. Instead of the electronic backplane, the OPPSCM is deployed. The complete system is shown in Fig. 2 schematically.

Each light-guiding plate of OPPSCM is covered by claddings and is optically isolated from the neighbour plates. The perimeters of the plates are formed by two circle segments and two straight lines, as shown in Fig. 3 (a) and Fig. 4. This structure was proposed by Y. Okada et al. in [3], but no light-guiding plates were used. It will be showed later that the plates offer several essential advantages.

The LDs and PDs are connected at the circumference of the large circle segment. The other borders of the plates formed by the small circle segment are put on a cylinder mirror with the same radius of the small circle. The optical axes of the LDs and PDs, respectively, point at the cylinder mirror by different angles φ (Fig. 4). φ depends on the position at the circle segment given by angle β. The relation between φ and β is shown in Fig. 5.

By means of the appropriate choice of the cylinder mirror radius and the plate geometry the optical power of the LDs impinging on the cylinder mirror is reflected back to the whole large circle segment. At an arbitrary position of the circumference of the large circle segment optical power can be detected by PDs. The angle ψ is assumed the power half angle of the emitting diodes (see Fig. 6).

3. Measurement and simulation results

The detected optical power by a PD at an arbitrary position determines essentially the bit error rate (BER) of the optical digital data transmission, if the optical receiver is already designed. At the backplane level a BER less than 10^{-12} is necessary. The optical receivers for fiber transmission are inappropriate for application in the optical backplane systems because of the high power dissipation. The optical receivers for the backplane level data transmission are still developed. Therefore we calculate and measure, respectively, here only the optical power which is detected by a PD with the active area of A.

Considering a LD at the position given by the angle β and a PD at the position γ (see figure 4), the optical power P_O impinging on the active PD area can be calculated by integration of radiation characteristic of the LD multiplied by the solid angle from φ_1 to φ_2 and from $-\theta_t$ to $+\theta_t$, if the active PD area is assumed to be rectangular:

$$P_O(\beta,\gamma) = \int_{-\theta_t}^{+\theta_t} \int_{\varphi_1(\beta)}^{\varphi_2(\beta)} R \cdot I(\theta,\varphi) \cdot \cos\theta \cdot d\theta \cdot d\varphi \qquad (1)$$

the angle θ_t can be obtained using Snell's law:

$$\theta_t = \arccos\left(\frac{n_2}{n_1}\right) \qquad (2)$$

φ_1 and φ_2 can be calculated numerically by computer simulation. In equations (1) and (2) R is the reflection factor of the cylinder mirror, $I(\theta,\varphi)$ is the radiation characteristic of the LD, n_2 is the refractive index of the cladding, n_1 is the refractive index of the plate core (Fig. 3 (b)). Because of the short distances of backplanes the absorption of the light-guiding plate is negligible.

In our measurement we applied a cylinder glass mirror with aluminium coating. The light-guiding plate consists of PMMA surround by air, i.e. $n_1 = 1.49$ and $n_2 = 1.0$. As emitting diode an AlGaAs infrared LED with the peak wavelength of 850 nm was used. An internal mounted spherical lens focused the emitted optical power so that the power half angle of the LED was about 26° which is similar to the power half angle of a LD. The radiation characteristic of the LED is rotationally symmetrical about its optical axis and is shown in Fig. 6. The driving current for the LED was about 30 mA. As receiving diode, a PIN PD with an active area of 1 mm^2 was used. The angular responsitivity of the PD was nearly a cosine function.

The calculation was carried out using the above mentiond parameters. The reflection factor of the cylinder mirror R in eq. (1) was assumed to be 1.

Figure 7 shows the calculated results (solid curve) compared with the measured results. Using an emitting diode which has a total optical power of about 1 mW a PD can detecte at least 3 μW. Without the PMMA plate the PD can only detecte 0.1 μW optical power at the same position as before. If a multi quantum well LD is used which emits totally 10 mW optical power at the driving current of 30 mA, a PD can detecte at least an optical power of 30 μW. At such power high speed optical receivers can be easily designed.

We also estimated the dispersion which is due to the different distances of the ray beam of the LED to a PD. Compared without the PMMA plate the dispersion increases by the total reflection at the cladding. But this increase will not considerably influence the transmisssion quality. Because the power half angle of a LD lies in the region about 20° or less, the dispersion is estimated to only several picoseconds. This is negligible for data transmission at backplane level.

4. Conclusion

Our first investigation results show that the optoelectronic backplane OPPSCM is very well suitable for board-to-board high-speed data transmissions. Besides nearly cross-talk free data transmission of every opto-electronic backplane, also at very high data rate, the following advantages of the OPPSCM are particularly emphasized:

- the structure of OPPSCM makes the application of LDs particularly easy, the angular characteristic of LDs is fully exploited without additional optical elements,
- a good separation of the optical channels at small pitches of LDs is possible by using the light-guiding plates, therefore compact systems can be built,
- the transmitted optical power increases up to 30 times compared to the power without light-guiding plates, using currently available LDs as emitting diode the optical power of 30 μW can be detected,
- the time delays from a LD to all PDs are almost equal, the clock skew due to the different delays on an electronic backplane can be reduced considerably.

References

[1] R. A. Nordin, et al., "A systems perspective on digital interconnection technology", IEEE J. of Lightw.Techn., vol. 10, June 1992, pp. 811-827

[2] J. Jiang, U. Kraemer, "Two new developments for optoelectronic bus systems", Proc. SPIE 1505, pp. 166-174

[3] Y. Okada, et al., "Dialog.H: A highly parallel processor based on optical common bus", COMPCON '83, IEEE Computer Society Press, 1983, pp. 461-467

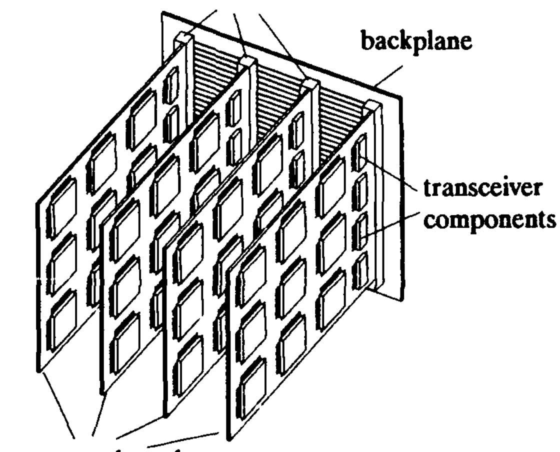


Fig. 1 Computer boards coupled to a backplane

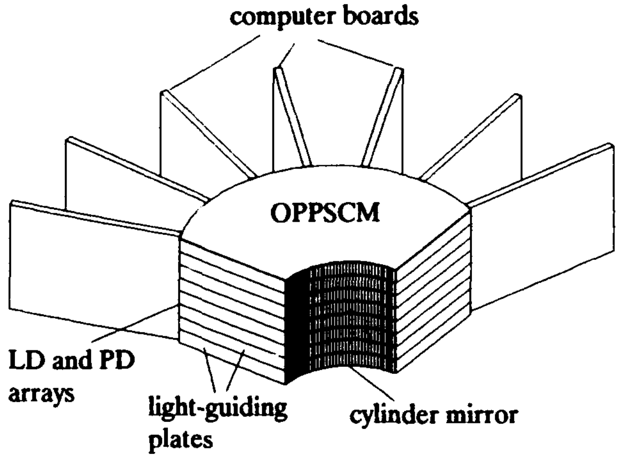


Fig. 2 A computer system with OPPSCM as backplane

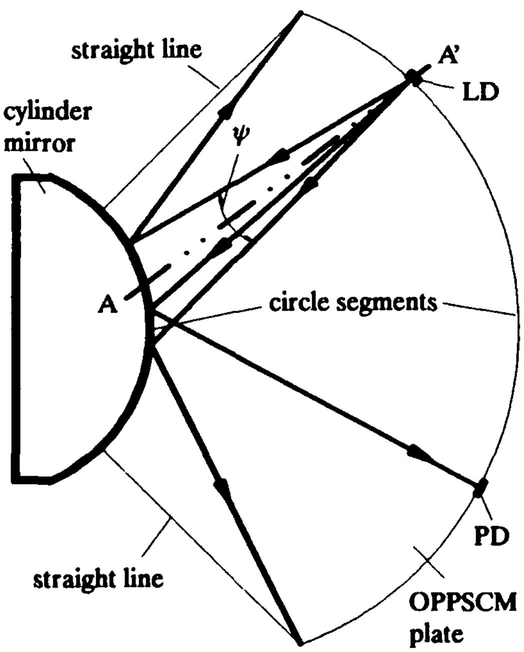


Fig. 3 (a) Top view of an OPPSCM plate

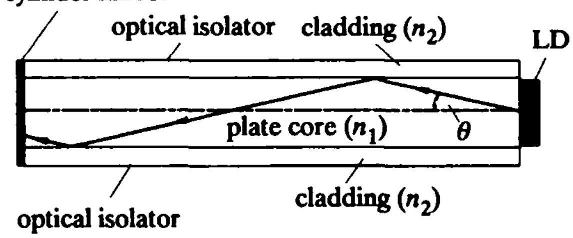


Fig. 3 (b) Side view of cut AA'
(different scale factor to (a))

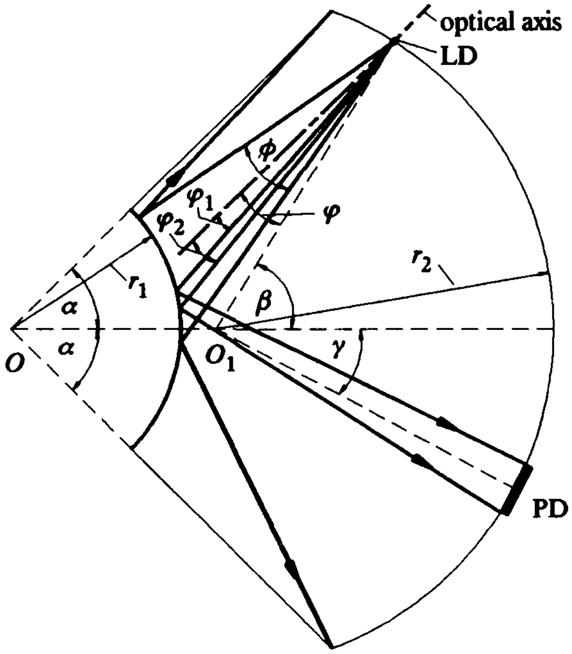


Fig. 4 Optical power transmission from a LD positioned at β to an arbitrary positioned PD at γ

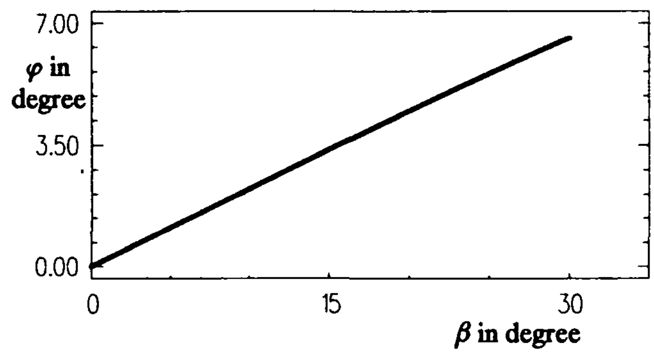


Fig. 5 The angle φ depends on the position given by β

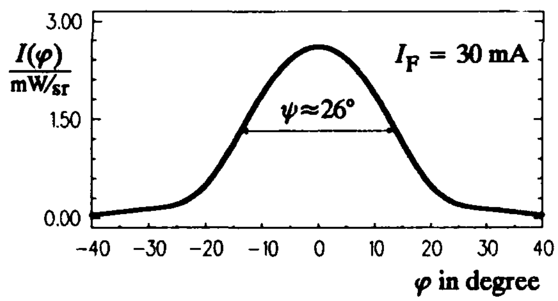


Fig. 6 Radiation characteristic of the LED used for the measurement

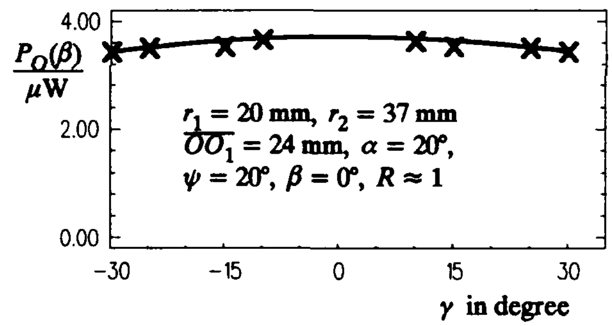


Fig. 7 (a) Calculated (solid curve) and measured (X) optical power depending on the PD position γ at LED position $\beta = 0°$

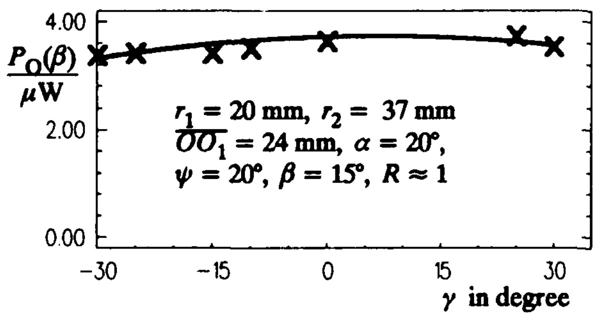


Fig. 7 (b) Calculated (solid curve) and measured (X) optical power depending on the PD position γ at LED position $\beta = 15°$

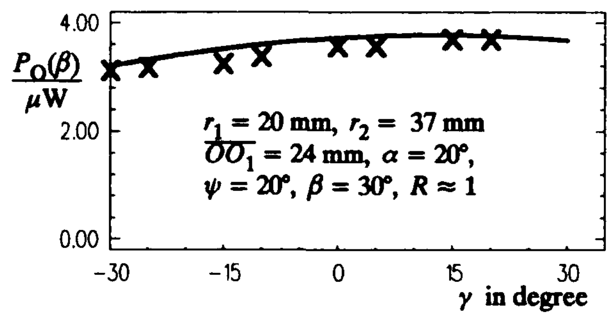


Fig. 7 (c) Calculated (solid curve) and measured (X) optical power depending on the PD position γ at LED position $\beta = 30°$

SYNCHRONIZING AND CONTROLLING FAST DIGITAL OPTICAL PROCESSORS

Vincent P. Heuring and Valentin N. Morozov
University of Colorado at Boulder
Optoelectronic Computing Systems Center†
Campus Box 425
Boulder CO 80309-0425

This paper discusses techniques for synchronizing and controlling fast digital optical and optoelectronic processors. When optical processors require significant amounts of control for proper operation, designers have traditionally resorted to electronic host computers. Such approaches are often inadequate because of high host computer cost, or the need for speed that an electronic host computer cannot deliver. We discuss how to design controllers in the optical or optoelectronic domain. Since data and control pulses must arrive in synchronization at all interaction points, system timing and resynchronization are discussed as the major problems to be solved in designing such systems. We present solutions to the problems of timing and resynchronizing high speed optical processors in the context of two system timing paradigms: the gate and strobe paradigm and the more recent time-of-flight paradigm. Gate and strobe designs are synchronized by gating the data pulses from storage element to storage element. Time of flight systems can be synchronized by *pulse reshaping*: resynchronization of the weakening pulse by retiming its leading and trailing edges and by *clock gating*: gating a fresh copy of the master clock signal to replace a weakening signal pulse.

1. INTRODUCTION

1.1. The Need for Control

Most of the optical and optoelectronic processors being designed and constructed posit electronic host computers not only as data sources and sinks, but also as processor controllers.[1, 2]. Guilfoyle,[3] has estimated that his DOC II optical processor will have an input data rate of 12.8 Gbits per second, and has observed that the control and data flow in the electronic domain will tax the limits of electronic design. The main focus in the design of these and similar systems has been on the flow of data through the processor, the means of processing, or the interconnections, not on the control of the system.

The control task is particularly difficult if arrays of processing elements require reconfiguration as in the designs of Murdocca and Huang.[4, 5] If each processing element must be reconfigured for each processing cycle, if N control bits are required per processing cycle, and there are M processing cycles, a controller with a memory of up to MxN bits is responsible for emitting an array of N bits once each cycle, repeated for M cycles. Control a 32x32 array of processing elements reconfigured at a 1 GHz rate would require 1 Terabit per second of control information. The timing of that bit-emission introduces another level of design complexity, which must be addressed and considered from the first step in the design process. Furthermore, if the processor wishes to lay any claim to programmability or to having any general-purpose nature, then the controller must incorporate the ability to loop and branch conditionally or unconditionally, adding yet more complexity.

1.2. Traditional Control Unit Design

In the electronic domain, these controllers are designed using either a hardwired approach, where random logic is used to implement the control signals, or a microcode approach, where each control word is stored

† This work was supported by the National Science Foundation ERC program, grant number CDR 8622236, and by CATI, the Colorado Advanced Technology Institute.

in a microcode control store to be retrieved as required. The microcoded controller begins a control sequence when it is presented with the address of the first control word. That control word contains not only the control bits but also the address of the next control word: a finite automaton. The microcode approach, which is generally to be preferred in optical and optoelectronic processors, requires a fast, wide-word ROM.

2. TIMING AND SYNCHRONIZATION OF FAST OPTICAL PROCESSES

2.1. Issues in Timing and Synchronization

Digital optical computers can also be viewed as finite state machines that compute the function: NextState = (PresentState, inputs). In simple terms this means the recirculation of information from storage (memory) elements to processing elements and back to storage elements, under programmed control. There are two important issues to be addressed in designing these systems: information must flow in a synchronized fashion to and from processing elements and memory elements, and in order for this synchronization to be maintained, the signal and control pulses must retain their shape throughout the process. For example, an array of data pulses emitted from some array of storage elements must be in synchronization when arrive at their respective processing elements and must remain in synchronization until they arrive at the next set of storage elements. Synchronization loss and pulse degradation can be caused by pulse dispersion, pulse jitter, long pulse rise or fall times of processing elements, loss of pulse amplitude or energy caused by loss in passive elements of the system, and *signal skew.* Signal skew is defined as the difference in arrival time between the earliest arriving pulse and the latest arriving pulse at a point of interaction. In spite of all these deleterious factors the pulses must arrive at their points of interaction in synchronization and with sufficient pulse strength to stimulate the proper response in their respective detectors. We have previously developed the formalism to permit the designer to consider synchronization and power loss and crosstalk during the system design phase[6, 7]. This paper considers the practical means by which the designer can restore pulse synchronization and pulse amplitude in the face of these deleterious effects.

2.2. The Gate and Strobe Paradigm

How the designer approaches the issues of maintaining synchronization and pulse shape depends on the nature of the processing and storage elements. If the pulse rise or fall time or signal skew is an appreciable fraction of the propagation time between or through processing elements, then the designer usually arranges for the pulse duration to be equal to the propagation time from storage elements to processing elements and back to storage elements. This is usually referred to as the *gate and strobe* method. Data pulses are *gated* from the storage elements into the processing elements and *strobed* into the next set of processing elements. The gating control pulse, applied to the first set of storage elements, is of sufficient duration to permit the data pulses to propagate to the next set of storage elements. The strobe control pulse, applied to the second set of storage elements, need only have sufficient duration to allow entry of the data pulses into the storage elements. Figure 1 shows the data and control flow in a gate and strobe system.

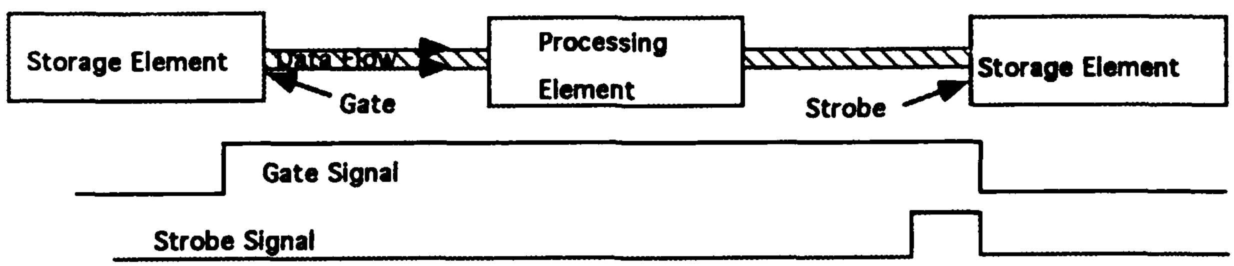


Figure 1: The Gate and Strobe Process for Controlling Timing and Synchronization

In a gate and strobe system, synchronization is ensured by propagating only a single, wide pulse through the circuit from storage element to storage element. Maintaining or restoring pulse amplitude or power depends on the nature of the devices in the implementation domain. If the devices are active, and inherently restore amplitude or power, then the designer only needs to ensure that fanout and loss are kept within the proscribed limits of the devices' ability to accept and transmit the signals. If the devices are passive then the amplitude or power must be restored by a restoring device – typically by ANDing the signal with a copy of the clock[8], often referred to as clock gating.

2.3. The Time-of-Flight Paradigm

If the pulses' rise and fall times and skews are short compared to the propagation time between devices, then the designer may chose to employ pulses that are small compared to propagation times. In this case, the designer also has the opportunity to resynchronize and restore amplitude or power, "on the fly" so to speak, without latching pulses into storage elements. That is, at a point in the system when the data pulses are dangerously out of synchronization or low in amplitude or power, the designer employs circuit elements specifically to resynchronize pulses and to restore their shape, without any slowing down or stopping (latching) of the pulses at storage elements. Information flows through the system at a speed governed strictly by propagation time through the system. This class of system is known as a "time of flight" architecture. These architectures are characterized, then, by pulses that are short with respect to propagation times, that are not latched into storage elements, but rather periodically restored in amplitude and phase as they flow through the circuit. One can view such systems as operating at their *natural,* or *resonant* frequencies.

Time of Flight systems are synchronized by controlling the propagation times of data and control pulses through the system. Such a task can be exceedingly complex if many feedback loops are used. The algorithms described in[6] are adequate to compute propagation times in single thread designs, but require extensions to parallel and matrix designs. Restoring pulse amplitude or power levels can be accomplished in the same manner as for gate and strobe architectures above, but there is the additional complication of needing to restore pulse *width* also: as pulses propagate through a time of flight system, they spread and lose amplitude. This means that the leading and trailing edges of the pulses are invalidated, and must be trimmed or excised before the resynchronization operation, resulting in a shortened pulse. Thus time of flight systems must include a mechanism for pulse stretching. This may be accomplished by splitting each pulse in two, and delaying one path with respect to the other, resulting in a pulse of longer duration and lower amplitude. At this point, the pulse may be ANDed with a synchronizing master clock signal, resulting in a fresh, resynchronized, amplitude-restored pulse. Figure 2 shows the data and control flow in a time of flight system.

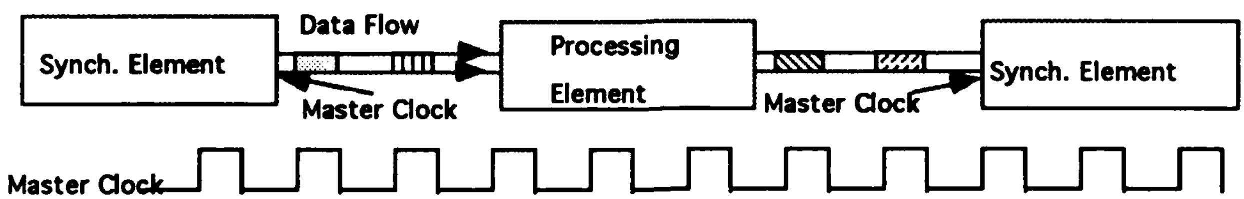


Figure 2: The Time of Flight Process for Controlling Timing and Synchronization

3. EXAMPLE OF AN OPTOELECTRONIC CONTROLLER

3.1. Controller Design

Figure 3 shows the design of an optical controller we are constructing to control a parallel optoelectronic processor. The implementation domain of the controller is optoelectronic NOR gates interconnected by free-space using holographic optical elements to control the geometry of information flow. Here pulses recirculate from the initial NOR gate array through a fanout hologram to another NOR array that selects a minterm. The gate that is selected illuminates a given hologram that contains control information that is "narrowcast" to processor elements to be controlled. Next address information is also contained in the hologram. That information is sent back to the first NOR gate array.

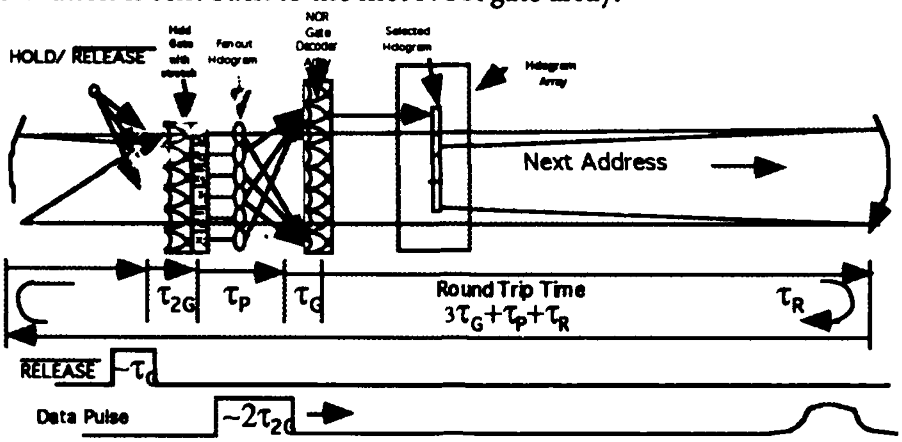


Figure 3. Optoelectronic control unit using NOR gates and holographic optical elements

Synchronization is achieved as follows: notice that when a NOR gate is illuminated, that gate is disabled. This is the function of the **RELEASE** signal. The **RELEASE** signal is held active (that is, on) as the invalid, leading-edge portions of the data pulses arrive. When the valid portion of the pulses arrives, the **RELEASE** is made low, thus "releasing" the data signals to the first NOR gate array. This gate array has the pulse stretching property designed into it by delaying a portion of the arriving signal, thus effectively stretching the pulse from an initial width of τ_g to a width of $2\tau_g$. Details of the controller's operation can be found in [9].

4. CONCLUSIONS

Controllers for optical and optoelectronic processors can be designed using the same implementation domain as the processor itself. The system can be synchronized using either the gate and strobe method or the time of flight method, depending on data pulse characteristics. System timing and synchronization must be considered from the earliest part of the design process.

5. REFERENCES

1. M.E. Prise, N. Stribel & M.M. Downs. *Optical and Quantum Electronics* "Optical considerations in the design of digital optical computers," 30, 49-77 (1988).
2. G.C. Marsden, A.V. Krishnamoorthy, S. Esener & S.H. Lee. *Optics Letters* "Dual-scale topology. Applied Optics optoelectronic processor," 16, 1970-1972 (1991).
3. P.S. Guilfoyle. *Digital Optical Computer II (DOC II): Performance Specifications* (Optical Society of America, Salt Lake City UT, 203-206 1991).
4. M.J. Murdocca. *A Digital Design Methodology for Optical Computing* (MIT Press, Cambridge MA, 1990).
5. A. Huang. *Applied Optics* "Computational Origami - The Folding of Circuits and Systems," Applied Optics, 31, #26, 5419-5422, (1992).
6. J.P. Pratt & V.P. Heuring. *Applied Optics* "Delay Synchronization in Time-of-Flight Optical Systems," 31, 2430-2437 (1992).
7. J.P. Pratt & V.P. Heuring. *Applied Optics* "Designing Digital Optical Computing Systems: Power Distribution and Crosstalk," 31, 4657-4661 (1992).
8. V.P. Heuring, H.F. Jordan & J.P. Pratt. *Applied Optics.* "A Bit Serial Architecture for Optical Computing," 31, 3213-3224 (1992).
9. V.P. Heuring & V. Morozov. *An optically controlled digital optical matrix processor, SPIE,* 1773b, paper 03, San Diego, CA, July 19-24, 1992.

Model of Lossless Bus Structure Using Erbium Fiber Amplifiers Pumped near 820nm

Manoj M. Bidnurkar, Steven P. Levitan, Rami G. Melhem, and Donald M. Chiarulli
University of Pittsburgh, Pittsburgh, PA 15260

In this paper we present a model and simulation results for a lossless tapped fiber bus using Erbium fiber amplifier segments at a pump wavelength near 820nm. Analytical models for fiber amplifiers were developed by Giles, et. al.[Gie:91] and by Sunak, et. al.[Sunak]. These results adapt the model developed by Giles, et. al.[Gie:91]. to bus applications. The adapted model breaks the fiber into small segments and then solves a pair of coupled differential equations for each segment. The process is iteratively repeated for each segment along the length of bus.

Between each amplifier segment in the bus model are one or more passive, 2×2, symmetric couplers. The specific placement of amplifier segments, and the gain characteristics of each are studied relative to the number and coupling ratio of taps between the segments. The results of the simulations show that, with a proper choice of the amplifier characteristics and placement strategy, it is possible to build a lossless optical buses In addition, non-linear gain characteristics shown by the model at low pump powers and short fiber lengths, suggest that that the logic one power is maintained without loss, while logic zero, powers show a net attenuation.

Amplifier Model

The light traveling in the fiber is assumed to be composed of a number of optical beams [Gie:91]. Each beam has a central frequency ν_k, and a frequency spread $\Delta\nu_k$ around the central frequency. The variable k is a dummy index which is summed over the total number of optical beams. Each of these beams, affects the populations of the Er-ions in the various energy levels.

The analysis is done in cylindrical co-ordinates r, ϕ and z. Here r is the distance in the direction perpendicular to the fiber axis, z is the distance along the fiber axis and ϕ is the azimuthal angle. The distance z is measured from the point at which the signal is introduced into the fiber. In the model the pump and the signal, are assumed to be co-directional.

The light intensity of the k^{th} optical beam at any point (r, ϕ, z) in the fiber is given by $I_k(r, \phi, z)$. Since, the intensity is power per unit area, the total power at a cross-sectional plane $P_k(z)$, is given by the integration of $I_k(r, \phi, z)$ over the cross-section.

The model makes certain assumptions regarding the light propagation in the fiber. Firstly, only two optical beams are taken into consideration. One for the pump $k = 1$ and the other for the signal $k = 2$. This means that we have only two coupled differential equations to be solved. The second assumption is that, the fiber is assumed to be a single mode fiber. This implies that only the zeroth modes are dominant for the pump and the signal.

$$\frac{dP_k}{dz} = \frac{\alpha_k + g_k^*}{\Gamma_k} P_k \int_0^{2\pi} \int_0^b \frac{n_2(r, \phi, z)}{n_t} i_k r \, dr \, d\phi - (\alpha_k + l_k) P_k$$

$$+\frac{g_k^*}{\Gamma_k} h\nu_k \Delta\nu_k \int_0^{2\pi} \int_0^b \frac{n_2(r,\phi,z)}{n_t} i_k r\ dr\ d\phi \tag{1}$$

The values for the variables used by the model are given in Table 1.

Table 1: Symbols and Values

Symbol	Value
σ_{a1}	$0.18 * 10^{-25} m^2$[Min:91]
σ_{a2}	$5.1 * 10^{-25} m^2$[Bar:91]
σ_{e1}	$0\ m^2$
σ_{e2}	$4.4 * 10^{-25} m^2$[Bar:91]
Γ_1	0.6[Bar:91]
Γ_2	0.6[Bar:91]
n_t	$10^{24} - 10^{26}$ per m^3[Min:91, Agg:91]
ν_1	$3.66 * 10^{14}$ per *sec*
ν_2	$1.94 * 10^{14}$ per *sec*
$\Delta\nu_1$	$4.462 * 10^{12}$ per *sec*
$\Delta\nu_2$	$1.249 * 10^{12}$ per *sec*
τ	10 *ms*[Bar:91]
b	6.2 *microns*

Let P_{k1}, P_{k2}, $P_{k3}, \ldots$, be the powers of the k^{th} optical beam, at the inputs of pieces 1, 2, 3,..., respectively. Since the pieces are end to end, P_{k2}, $P_{k3}, \ldots$, are also the powers at the outputs of pieces 1, 2,..., respectively. Let $z_1(=0)$, $z_2, z_3, \ldots$, be the distances of the inputs of the pieces from the beginning of the fiber. For a small piece of fiber, Equation 1 can be approximated by:

$$\begin{aligned} dP_k &= (\frac{\alpha_k + g_k^*}{\Gamma_k} P_k \int_0^{2\pi} \int_0^b \frac{n_2(r,\phi,z)}{n_t} i_k r\ dr\ d\phi - (\alpha_k + l_k) P_k \\ &+ \frac{g_k^*}{\Gamma_k} h\nu_k \Delta\nu_k \int_0^{2\pi} \int_0^b \frac{n_2(r,\phi,z)}{n_t} i_k r\ dr\ d\phi) dz \end{aligned} \tag{2}$$

where dP_k is the difference in the powers at the input and the output of the piece and dz is the length of the piece. Thus for the first piece we can write

$$\begin{aligned} P_{k2} - P_{k1} &= (\frac{\alpha_k + g_k^*}{\Gamma_k} P_{k1} \int_0^{2\pi} \int_0^b \frac{n_2(r,\phi,z)}{n_t} i_k r\ dr\ d\phi - (\alpha_k + l_k) P_{k1} \\ &+ \frac{g_k^*}{\Gamma_k} h\nu_k \Delta\nu_k \int_0^{2\pi} \int_0^b \frac{n_2(r,\phi,z)}{n_t} i_k r\ dr\ d\phi)(z_2 - z_1) \end{aligned} \tag{3}$$

Integration of the Amplifier and the Bus Models

The linear bus consists of a single fiber running from the input to the output. The data signal can be tapped and new data or pump introduced at any coupler in the bus. The tapped signal is fed to a detector. Each coupler attenuates the signal due to fixed losses due to reflection and excess loss, and by the power transferred to the tap fiber. The power transfer for each wavelength is given by the coupling ratio of the ratio. The models for the linear bus and the fiber amplifier are integrated in figure 1. In this bus, an Erbium amplifier can be introduced between any two couplers in the bus. Otherwise the two couples can be connected by an ordinary fiber.

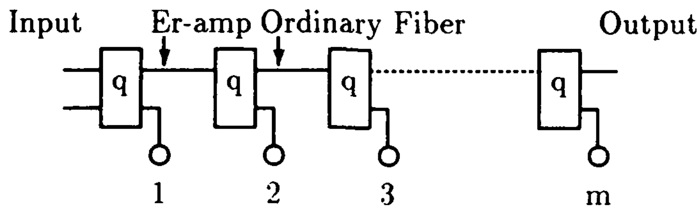


Figure 1: Linear Bus With Er Amplifiers

Simulations

Figure 2 shows the gain characteristics for amplifier segment modeled. By the nature of this application, a small segment of fiber, $1meter$, with Er-ion density of $2.5 * 10^24/m^3$ was modelled. This approach provides for minimum latency, while matching the relatively small gain to the losses anticipated the tap segments. For this plot, the input data signal was varied from $10\mu W$ to $100\mu W$, in steps of $10\mu W$. Each curve in the plot represents an input pump power, varied from $10mW$ to $50mW$, in steps of $10mW$.

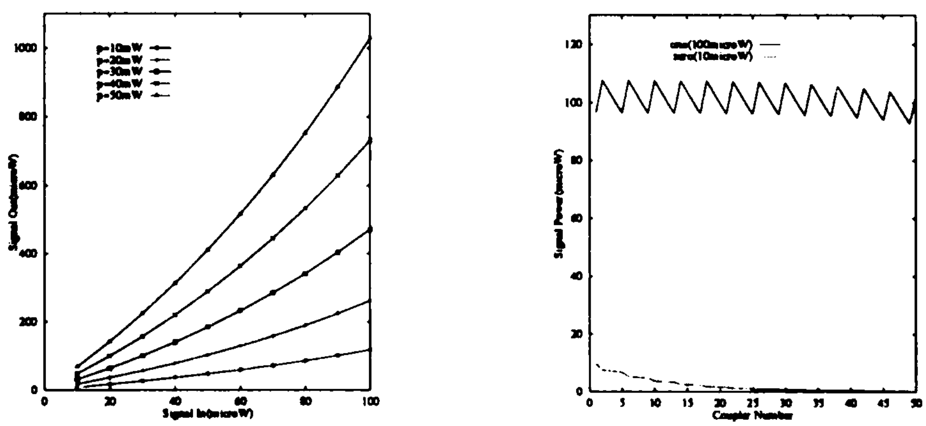


Figure 2: a) Amplifier Input/Output b) Logic levels on 50 tap bus

The most significant feature of this curves is the difference in amplification for as signal power is increased. For example at a pump power of $10mW$, a logic zero signal in the range of $10\mu W$ show only a nominal gain. But, a signal of power of $100\mu W$, a typical logic one is significantly amplified. Thus by matching the gain factor from the logic one signal to to losses in the tap segment, this signal show no loss over the length of the fiber. However, the logic zero signal with its smaller gain will show a net attenuation.

This effect is shown in by the simulation documented by figure 2. In this case, amplifier segments of $1m$, pumped at $10mW$, were placed between groups for four passive couplers of ratio 0.964 for the signal wavelength and 0.01 for the pump wavelength. The bus tested had a length of 50 couplers.

From Figure 2, it can be seen that the logical one power level is approximately maintained through the bus while the logical zero dies out. Beyond about 40 couplers, the logical one starts falling. This is due to an inexact match between the signal gain in the amplifier and losses in the taps. Such minor differences can be easily compensated for by ocassionaly adjusting the number of taps/segment or using a somehat larger pump power in certain segment.

Acknowledgment

This research has been supported in part by grants from the Air Force Office of Scientific Research, and the National Science Foundation.

1 References

[Gie:91] Giles, Randy C. and Desurvire, Emmnuel, Modelling Erbium Doped Fiber Amplifiers. *Journal of Lightwave Technology*, Vol. 9, No.2, (February, 1991), pp. 271-283.

[Raj:90] Liu, Karen and Ramaswami, Rajiv, *Analysis of Optical Bus Networks Using Doped-Fiber Amplifiers*, Yorktown Heights, New York: IBM Research Division, June, 1990.

[Sunak] Bastien, Steven P., Sunak, Harish R. D., Balakrishnan, Sridhar and Kalomiris, Vas, E., *Analysis of Erbium Doped Fiber Amplifiers* Kingston, Rhode Island: University of Rhode Island.

[Lee:86] Lee, Donald L., *Electromagnetic Principles of Integrated Optics* New York: John Wiley & Sons, 1986, pp. 283-288.

[Min:91] Miniscalco, William J., Erbium-Doped Glasses for Fiber Amplifiers at 1550 nm, *Journal of Lightwave Technology*, Vol. 9, No. 2, (February,1991), pp. 234-250.

[Bar:91] Barnes, William J., Laming, Richard I., Tarbox, Eleanor J. and Morkel, P. R., Absorption and Emission Cross Section of Er^{3+} Doped Silica Fibers, *IEEE Journal of Quantum Electronics*, Vol. 27, No. 4, (April, 1991), pp. 1004-1010.

[Agg:91] Poulain, M., Fluoride Glass Composition and Processing, in *Fluoride Glass Fiber Optics*, Aggarwal, Ishwar D., Lu, Grant, ed., San Diego, Academic Press, 1991.

Simulated Annealing Applied to Placement of Processing Elements in Optoelectronic MultiChip Modules Interconnected by Computer Generated Holograms

D. Zaleta, J. Fan, C.K. Cheng, S.H. Lee
University of California, San Diego
Electrical and Computer Engineering Department
and Computer Science Department, Mail Code 0014
9500 Gilman Drive
La Jolla, CA 92093-0407
(619)-534-1743

1. Introduction and Background

New computer-aided tools are needed for the development of OptoElectronic MultiChip Modules (OE-MCMs) utilizing free-space optical interconnects. In the electronic world, one of the challenges in designing large systems is the physical placement and routing of electronic modules based on the system's physical and performance constraints. Such standard algorithms as Timber-Wolf [1] are commonly used to both place and route an electronic system given the netlist connecting the system's modules. Briefly, the placement problem attempts to find a mapping from a set of logical (opto)electronic processing elements (PEs) onto the physical position in the system, subject to the constraint of a list specifying the interconnections between the modules (commonly referred to as a netlist). Unfortunately, the standard algorithms used by electronics fail to properly model optoelectronic constraints. For instance, they minimize a cost function that incorporates the sum of all the interconnection distances and allow for a configuration space that allows the modules being placed to overlap. In the optoelectronic case, it has been shown[2] that for free-space optical interconnection of OE-MCM, it is the maximum interconnection distance that should be minimized and overlaps are not allowed. Recently, placement algorithms for optoelectronic systems using a matching algorithm have been developed[2,3]. In this paper, we utilize an adapted version of the simulated annealing algorithm to tackle the problem of placement and compare results for a twin butterfly architecture design example with the matching algorithm and a "straight-forward" placement.

2. Optoelectronic Placement Problem Formulation

Figure 1 shows the CGH interconnection model, we are assuming in this paper. It can be shown that the minimum features size, Δr, for an off-axis multi-phase level spherical diffractive lens is a monotonic decreasing function of the angle θ as shown in Fig. 2, with $f = 2$mm, $t = 1$cm, D=250μm, $\lambda = 0.8$μm, and m=2. Figure 2 shows that given a fixed fabrication limit, we are constrained to a fixed maximum bending angle, θ_m. Since $\tan(\theta_m) = d/t$ is fixed, the volume of the system (controlled by the parameter t) can be minimized by reducing the system's maximum d. As the number of PEs increase and the fanout from (and to) each PE increases and becomes irregular, the maximum d becomes increasingly difficult to minimize. Note that fanout can be accomplished by either segmented holograms of the type in Figure 1 or Dammann-type gratings combined with a quadratic phase function.

Figures 3 shows a schematic diagram of the physical models for a multi-stage OE-MCM. In this physical model, we assume there exist N (for an N-1 stage network) PE planes, P_k, that are interconnected to adjacent PE plane(s) by multiple optical links. Each link is considered to be of the configuration shown in Fig. 1 and thus the CGH planes are arrays of off-axis diffractive lenses. We assume that each link begins at a source (modulator or laser diode) and terminates at a detector. Note that this model as well as the algorithm itself is not limited to OE-MCM configurations but could be easily used also for wafer scale integration.

Figure 4 shows the logical diagram of the placement problem that we use to formulate a discrete configuration space that simulated annealing can handle. We assume that there exist a netlist which determines the interconnection pattern between a set of logical PEs (e_{kr}) in set E_k and the sets E_{k-1} and E_{k+1}. The assignment (or placement) of these logical PEs into the actual physical array slots, p_{ki} of plane P_k by the mapping σ_k determine both the interconnection distance (which determines the minimum volume and CGH synthesis) and direction (which determines the CGH synthesis).

3. Simulated Annealing Algorithm Applied to the Placement Problem

Simulated annealing is a well-known algorithm developed by Kirkpatrick, et. al. [4] to solve problems that commonly occur in combinatorial optimization. In order to use simulated annealing, the problem must first be configured as a system with a discrete number of states. The suitability of any state is given by a cost function, in our case C_k, that describes some feature about the system that the algorithm is to minimize. Given some initial state, the system is incrementally changed to a neighboring state, at which point a change in the cost function, ΔC_k, is calculated. If this change is negative, we adopt the new state. If it is positive, the new state is accepted if a random number (chosen uniformly between 0 and 1) is greater than $\exp(-\Delta C/T)$, where T is a constant referred to as the temperature. This process continues until the algorithm can no longer make any improvements. It is this probabilistic acceptance of new states that increase the cost, that results in the "hill-climbing" capability that prevents simulated annealing from getting caught in a local minimum. At the beginning of the algorithm, the temperature is set very high, so that almost all states are accepted. Periodically as the state evolves, the temperature is reduced so that fewer and fewer worse states are accepted. Finally ending up as a greedy algorithm, where only changes that result in a lower cost function are accepted.

The simulated annealing algorithm employed in the placement problem is shown in Fig. 5. The states are described by a given placement of each of the logical PEs e_{kr} to one of the physical locations on its respective plane, p_{ki}. An incremental change in a given state is defined by switching the placement of two random PEs on a given plane and then measuring the change in the maximum interconnection distance over that plane.

For the purposes of minimizing the system's volume, we are interested in the interconnect distance d_{kimj} between a PE, p_{ki}, on PE plane P_k and an interconnected PE p_{mj} on PE plane P_m given by

$$d_{kimj} = d\left(p_{ki}, p_{mj}\right) = \sqrt{\left(x_{ki} - x_{mj}\right)^2 + \left(y_{ki} - y_{mj}\right)^2} \tag{3.1}$$

where (x_{ki}, y_{ki}) and (x_{mj}, y_{mj}) are the coordinates of the two PEs (p_{ki} and p_{mj}, respectively). We define the cost function for the simulated annealing algorithm to be based on the maximum interconnection distance from each plane (say P_k) to any adjacent planes (P_{k-1} and P_{k+1}),

$$C_k = \max_{ij}\left\{d_{ki(k-1)j}, d_{ki(k+1)j}\right\} \tag{3.2}$$

The goal of the algorithm is then to minimize this cost function over all stages, $k, l = 1, 2, .., N-1$. Note that if p_{ki} and p_{mj} are not connected by the netlist then d_{kimj} for that pair is not defined.

The algorithm starts with a beginning configuration state (an initial placement for all PEs) from which we switch the placement of two PEs. One of the PEs to be switched represents the PE with the largest interconnect distance for that plane while the other PE is randomly chosen (with a uniform distribution) from the remaining PEs. The change in the cost function for the new state is calculated and it is determined if the new state is accepted according to the method described above. The algorithm then proceeds to the next plane. After all the planes have been altered a given number of times the temperature is lowered (exponential decay) and the process continues until no changes are accepted or the system cannot be reduced further.

4. Placement Results of the Simulated Annealing Algorithm

Simulated annealing was applied to the placement problem for a real twin butterfly design example. To serve as a benchmark, we compared the results of the algorithm with a straight-forward placement, i.e. raster order placement of all PEs into the PE planes, and the matching algorithm presented in Ref. 2. The straight-forward placement was used since it yields good results for the standard butterfly relative to any other systematic placements that we tried. The best simulated annealing result was 4.24 while that of the matching algorithm and straight-forward placement were 4.24 and 8.60, respectively (1 being the interconnection distance between a PE and his nearest neighbor). Figure 6 shows a histogram of the interconnection distances over the entire system. From these placement results, a 1st level mask was generated and statistics over all the interconnections were gathered comparing the algorithms with this standard placement. The results show us that the straight-forward placement has a small number of CGHs with large interconnection distances that result in an increase in the system volume while the algorithms reduced the maximum interconnection distance but increased the overall sum of the distances (or CGH complexity). Both the simulated annealing algorithm and the matching algorithm improved the system volume by ~50% over the straight-forward placement.

5. Conclusions

We have shown that due to the fabrication limitations of CGHs, highly parallel processing systems based on OE-MCM require new placement algorithms to reduce the system volume and to synthesize the CGH array. We have formulated the placement problem as the attempt to find the optimum placements for the logical PEs onto the physical PE slots based on a given netlist so that the maximum interconnection distance over all the stages is minimized. A simulated annealing algorithm was applied to the problem and compared to the matching algorithm and the straight-forward placement. The two algorithms resulted in reducing the overall system volume by 50% over a straight-forward placement. Based on these results, CGHs have been designed and fabricated to implement the optical interconnections[5].

6. References

1. C. Sechen and A. Sangiovanni-Vincentelli, "The Timber-Wolf Placement and Routing Package", *Proc. Cust. Int. Cir. Conf*, (1984).
2. J. Fan, D. Zaleta, C.K. Cheng, S.H. Lee, "Physical layout for computer generated holograms for optoelectronic multichip modules", *Submitted to IEEE MCM Conference*, (1992).
3. J. Fan, D. Zaleta, C.K. Cheng, S.H. Lee, "Physical synthesis of computer generated holograms in optoelectronic systems", *Submitted to DAC '93 Conf.* (1992).
4. C. Gelatt S. Kirkpatrick and M. Vecchi, "Optimization by Simulated Annealing", *Science*, **220**, 671-680 (1983).
5. B. Kress and S.H. Lee, "Iterative Design of CGH for Free-Space Optical Interconnections", *Submitted to OSA 1993 Topical meeting on Optical Computing*, (1992).

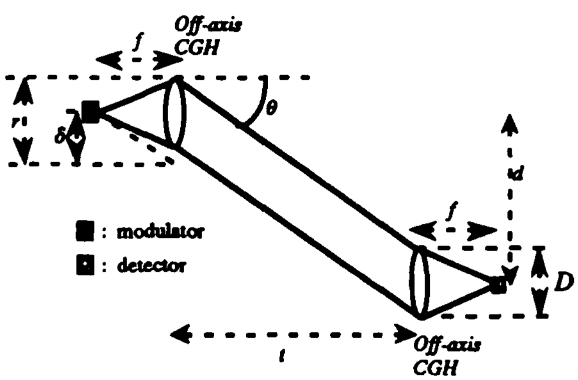


Figure 1. Schematic of CGH interconnect

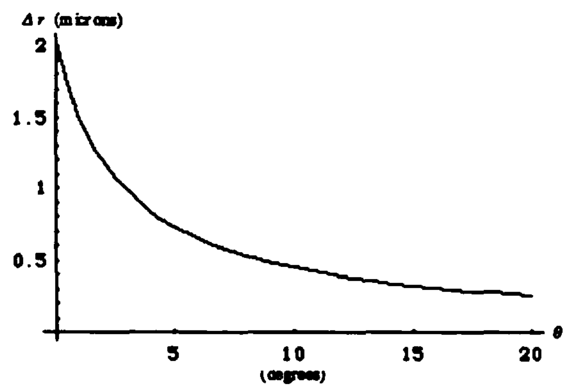


Figure 2. Relationship between the minimum features size, Δr, and the angle, θ.

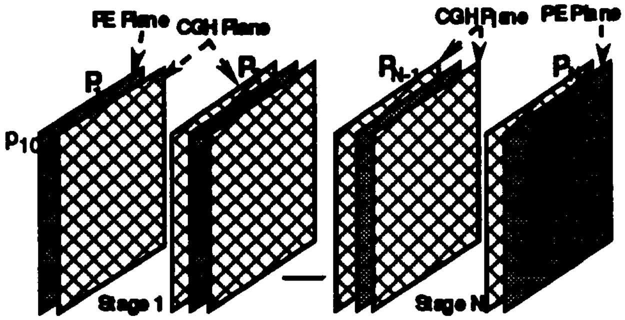


Figure 3. OE-MCM physical model

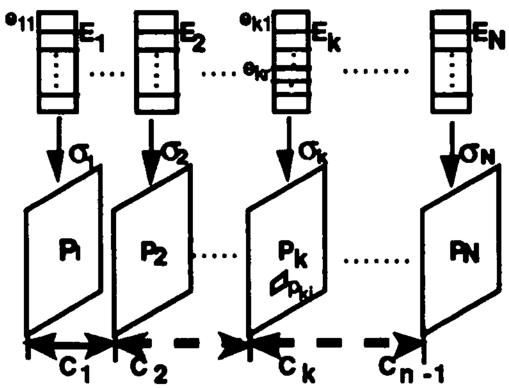
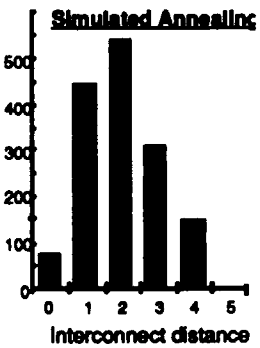


Figure 4. Logical relationship inside a general multistage interconnection network system.



(a)

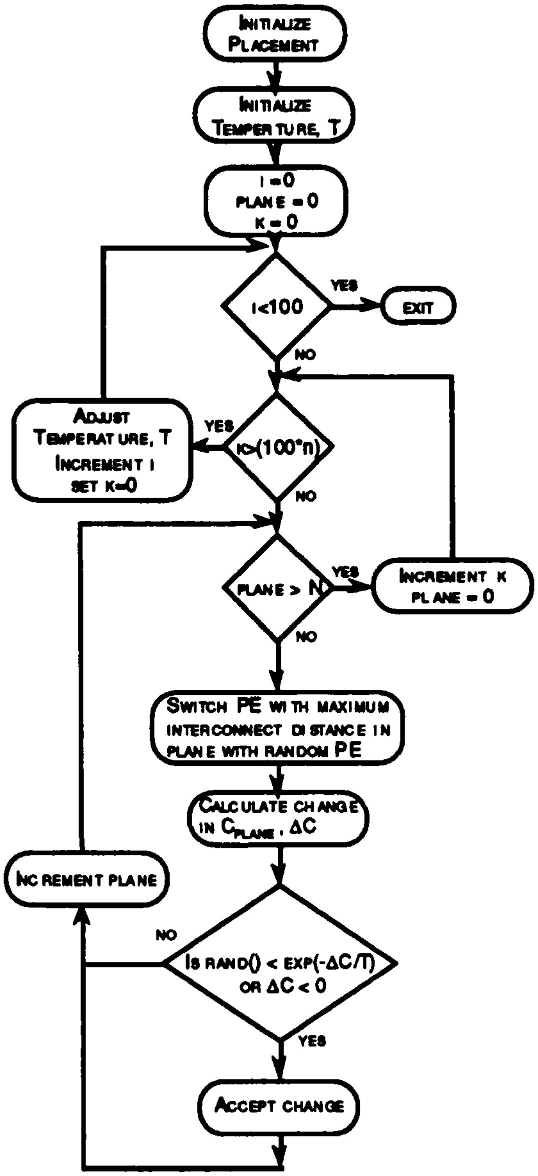
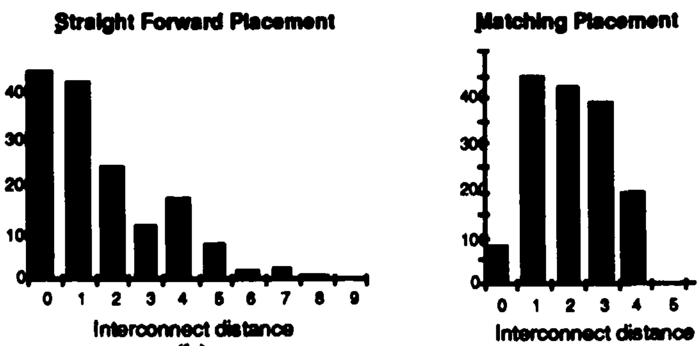


Figure 5. Flow chart of the simulated annealing algorithm applied to the placement problem.



(b) (c)

Figure 6. Histogram of the interconnect distances for the placement of a 64 node, 6-stage twin butterfly using (a) simulated annealing placement, (b) straight-forward placement, and (c) matching algorithm placement.

One-step Modified Signed-Digit Addition / Subtraction Based On Redundant Bit Representation

Hongxin Huang, Masahide Itoh and Toyohiko Yatagai
Institute of Applied Physics, University of Tsukuba. Tsukuba, Ibaraki 305, Japan

1 INTRODUCTION

Full parallel processing requires the parallel algorithm technique, in which consistes of a suitable number systems and an efficient encoding scheme for handing the data, and parallel architectures for realizing the algorithm. The modified signed-digit number system[1] has been proved to be able realize carry-free addition and subtraction, and its optical system is of potential two-dimension parallel processing ability.

Recently, optical implementation of MSD arithmetic by using the optical symbolic substitution has been proposed, and a number of schemes such as location-addressable memory (LAM), content-addressable memory (CAM), multichannel correlators and optical shadow-casting, have been suggested to implement optically symbolic substitution. The CAM method[2,3] is considered as an efficient method for performing MSD operations, but is suffer from the disadvantage that the number of reference pattern is proportional to the length of operands. The number of references becomes immense when performing high accuracy calculation. For example, for two 32-bit MSD addition, one require the storage of 1718 minterms with the one-step scheme proposed by Mirsaledi & Gaylord[2], and requires the storage of 700 minterms with the two-step scheme proposed by Li & Eichmann [3].

In this paper, we describe a new scheme of performing MSD addition / subtraction (A/S) in one step. The sixty-eight minterms (fourteen, twenty and thirty-four minterms for output digit 1, $\bar{1}$ and 0 respectively) for one-step MSD A/S are developed based on a new logic minimizing technique and the redundant bit representation of MSD number. Two new optical implementation methods, sampled discrete correlation and matrix multiplication, are proposed for implementing optically the one-step MSD A/S.

2 REDUNDANT BIT REPRESENTATION

Redundant bit (RB)representation, in which each bit of a digital number is encoded with several bits, has be proposed to reduce the symbolic substitution rules in digital optical computing[4]. One of the advantages of RB representation is that indescribable bit (number) can be expressed mathematically under given definition. For example, in MSD arithmetic, by using three-for-one RB codes that is digits $\bar{1}$, 0 and 1 are encoded with 001, 010 and 100, which can be denoted by [1], [2], and [4] respectively, the digit either $\bar{1}$ or 0 can be expressed by 011 or denoted by [5]. The physical means of this expression is that for performing the bit wise AND operation, the sum of all bits of the result is always 1. Similarly, either $\bar{1}$ or 1, either 0 or 1 and either $\bar{1}$ or 0 or 1 can be expressed with 101, 110 and 111, or denoted with [5], [6] and [7] respectively.

3 MINTERMS FOR ONE-STEP MSD ADDITION

Binary number system, the carry can be propagation from the least signification bit (LSB) to the most signification bit (MSB). The MSD number system limits the carry propagation to only two position to the left, hence, addition / subtraction of two MSD number can be carried out in three steps in a constants time regardless of the number of digit in MSD numbers. Each digit in the result, therefore, dependent on the three positions digit pairs. In a two-step scheme[3,5], two pairs are employed in the first step, and the affection of the third pair is done in the second step. In one-step scheme, all the three bit pairs have to be considered simultaneously. Because six trinary input digits are of 729 possible combination, in principle, 729 computation rules need be performed simultaneously for full parallel MSD additor (subtractor). Recently 729 computation rules is reduced to 56 minterms by using a logic minimizing technique[6]. The minterm has been reduced to thirty-four

with the new logic minimizing technique proposed and by utilizing RB representation of MSD digit. Table I lists the reduced minterms for MSD addition. Here the minterm is denoted with a bracketed string of six digits. This reduction is based on the following three facts:
(1) In same case, one can get output digits only examining two pairs digits. With this consideration, 369 computation rules instant of 729 computation rules is enough for one-step addition.
(2) Two computation rules can be combined into one if and only if the two rules map to the same digit and among three pairs digits there are five digits are same. With this consideration, 68 minterms (14, 20 and 34 minterms for output digit 1, $\bar{1}$ and 0 respectively) are found from the 369 computation rules.
(3) The output 0 can be considered as neither 1 nor $\bar{1}$, thus 34 minterms (14 and 20 minterms for output 1 and $\bar{1}$ respectively) is enough for performing one-step MSD addition.

The minterms for MSD subtraction shown in Table II is obtained with similar technique.

4 OPTICAL ENCODING

The optical implementation, both its algorithm and architecture, the complexity for performing one-step MSD A/S are strongly dependent on the encoding scheme employed. From Table I and Table II, it is clear that the minterm contains seven independent items, [1], [2], [4], [3], [5], [6] and [7]. An efficient encoding scheme is that it not only can represent three elementary terms [1], [2] and [4], which are corresponding to MSD digits $\bar{1}$, 0, and 1 respectively, but also can easily represent and distinguish other four complex terms, which correspond to either $\bar{1}$ or 0, either $\bar{1}$ or 1, either 0 or 1 and don't care bit (either $\bar{1}$ or 0 or 1). Figure 1 shows a typical encoding schemes, which use the same SBWP with that of triple-rail encoding of MSD digits, used in the following discussion.

5 OPTICAL IMPLEMENTATION WITH SAMPLED DISCRETE CORRELATION

In this scheme, each MSD digit is encoded by using three-for-one RB representation, which can be arranged as either a row vector or a column vector. When using column vector form, the two n-bits input operands as well as three padded bits are encoded into a $6\times(n+3)$ binary matrix (Fig.2), and each minterm is encoded as a binary matrix of dimension 6×3 . The thirty-four minterms are arranged side by side in the vertical direction, that is the memory matrix is matrix of 204×3, and the first ten minterms are minterms for output digit 1, others are for output digit $\bar{1}$. The discrete correlation result is a matrix of $209\times(n+5)$, a matrix of $34\times(n+1)$ can be obtained by sampling the correlation result. A binary matrix, which is of the characteristic that there is at last only one nonzero component in any column vector, is obtained after thresholding. The ith bit of the final result is determined by the ith column vector of the $34\times(n+1)$ thresholded matrix: if the ith column vector is a zero vector, the ith output bit is 0; if the nonzero component is one of the first ten components, the output is 1, else the output is $\bar{1}$.

The optical configuration for implementing this scheme is shown in Fig.3. Mainly, the setup is build up with three parts: an optical correlator constructed with two spherical lenses, a nonlinear threshold array (NTA) and a post-processor constructed with the cylindrical lenses. A LED array is used as two dimension incoherent source, the sampling processing is realized by sampling directly the light source by arranging properly the LED.

In practice, a strip-like one-dimension arranged memory mask can be rearranged in a square-like two-dimension pattern, so that the make the aperture of optical system minimum. Figure 4 shows the memory mask recoded thirty-four minterms, which is of aperture size of $(21d+6dx)\times(30d+4dy)$, where dxd is the size of cell for representation a pixel, dx and dy are the space between two adjacent minterms. The dx and dy are selected to avoid crosstalk between correlations produced by adjacent two minterms, and are of relation with the length of two operands.

6 OPTICAL IMPLEMENTATION WITH MATRIX MULTIPLICATION

In this scheme, the RB encoded minterm is arranged into a 1×18 row vector, thirty-four minterms is arranged into a 34×18 memory matrix as shown in Fig.5. The encoded input matrix for two n-bits operands ˙: of dimension of $18 \times (n+1)$ and is prepared as shown in Fig.6. Clearly the multiplication of the memory matrix with the input matrix is a matrix of $34 \times (n+1)$, which is the same as that obtained with the sampled discrete correlation scheme as described in above section. The similar threshold and post-processing, which can be performed with either optics or electric, are wanted for finding final result.

7 CONCUSSION

A new scheme, which employ 34 minterms, for one-step MSD addition / subtraction has been realized based on RB representation of MSD digits. Two new configurations, sampled discrete correlation and matrix multiplication, are proposed. Both methods use a fixed storage memory for any length operands and take full advantage of the parallelism of MSD number representation and along with the parallelism of optics.

REFERENCES

1. A.Avizinis, IRE Trans. on Electron Computers, Vol EC-10,389-400(1961).
2. M.M.Misalehi and T.K.Gaylord, Appl. Opt.25, 2277-2283(1986).
3. Y.Li and G.Eichmann, Appl. Opt. 26,2328-2333(1987).
4. Hongxin Huang and Toyohiko Yatagai, submitted.
5. N.Takagai, IEEE Trans on Compter, Vol C-34(9)789-796(1985).
6. M.M.Misalehi and T.K.Gaylord, Appl. Opt.25, 3078-3088(1986).

Table I Reduced minterms for one-step MSD addition

1	$\bar{1}$
[5 5 4 4 7 7]	[5 5 1 1 7 7]
[2 2 4 4 7 7]	[2 2 1 1 7 7]
[5 2 1 4 7 7]	[5 2 2 2 7 7]
[2 5 1 4 7 7]	[2 5 2 2 7 7]
[5 2 4 1 7 7]	[5 2 2 4 1 7]
[2 5 4 1 7 7]	[2 5 2 4 1 7]
[5 5 2 4 6 6]	[5 2 2 4 6 1]
[2 2 2 4 6 6]	[2 5 2 4 6 1]
[5 5 4 2 6 6]	[5 2 4 2 1 7]
[2 2 4 2 6 6]	[2 5 4 2 1 7]
[5 2 1 2 6 6]	[5 2 4 2 6 1]
[2 5 1 2 6 6]	[2 5 4 2 6 1]
[5 2 2 1 6 6]	[5 5 1 2 1 7]
[2 5 2 1 6 6]	[2 2 1 2 1 7]
	[5 5 1 2 6 1]
	[2 2 1 2 6 1]
	[5 5 2 1 1 7]
	[2 2 2 1 1 7]
	[5 5 2 1 6 1]
	[2 2 2 1 6 1]

Table II Reduced minterms for one-step MSD subtraction

1	$\bar{1}$
[5 5 4 1 7 7]	[5 5 1 4 7 7]
[2 2 4 1 7 7]	[2 2 1 4 7 7]
[5 2 1 1 7 7]	[5 2 2 2 7 7]
[2 5 1 1 7 7]	[2 5 2 2 7 7]
[5 2 4 4 7 7]	[5 2 2 1 1 7]
[2 5 4 4 7 7]	[2 5 2 1 1 7]
[5 5 2 1 6 3]	[5 2 2 1 6 4]
[2 2 2 1 6 3]	[2 5 2 1 6 4]
[5 5 4 2 6 3]	[5 2 4 2 1 7]
[2 2 4 2 6 3]	[2 5 4 2 1 7]
[5 2 1 2 6 3]	[5 2 4 2 6 4]
[2 5 1 2 6 3]	[2 5 4 2 6 4]
[5 2 2 4 6 3]	[5 5 1 2 1 7]
[2 5 2 4 6 3]	[2 2 1 2 1 7]
	[5 5 1 2 6 4]
	[2 2 1 2 6 4]
	[5 5 2 4 1 7]
	[2 2 2 4 1 7]
	[5 5 2 4 6 4]
	[2 2 2 4 6 4]

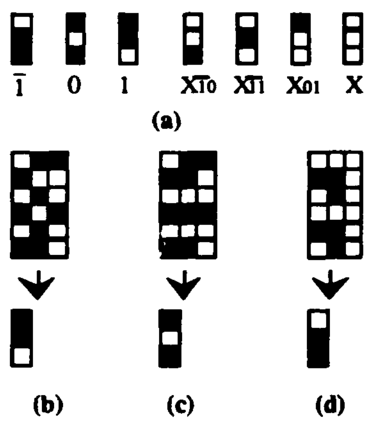


Figure 1 (a) Optical encoding of indepentent items of minterms for MSD addition / suntraction. (b), (c) and (d) Encoded computation rules for output digits 1, 0 and -1, which show the minterms [522166], [524266] and [551177] respectively of MSD addition

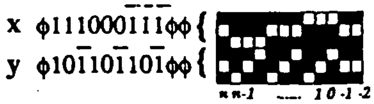


Fig. 2 Optical encode of input operands for sampled discrete correlation implemention. Two n-bits operands are encoded into 6x(n+3) binary pattern. Here ϕ denotes padded zero.

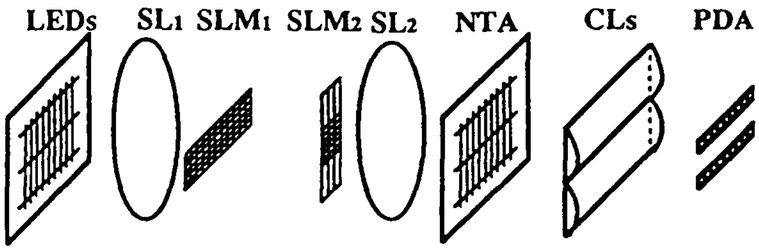


Fig. 3 Principle configuration for optical implementation of MSD arithmetic based on sampled discrete correlation. The LEDs are located at the cross points and the threshold elements are at the conjunction points .f LEDs

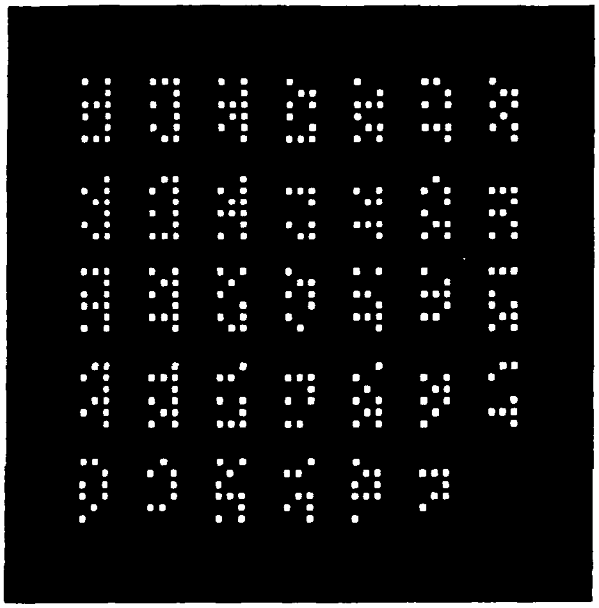


Fig. 4 Memory mask of 34 minterms for performing MSD addition with sampled discrete correlation

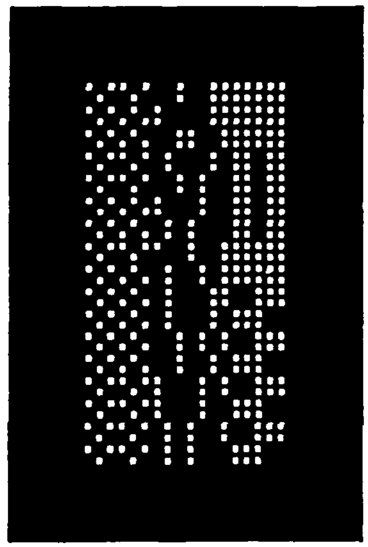


Fig. 5 Memory mask of 34 minterms for performing MSD addition with matrix multiplication.

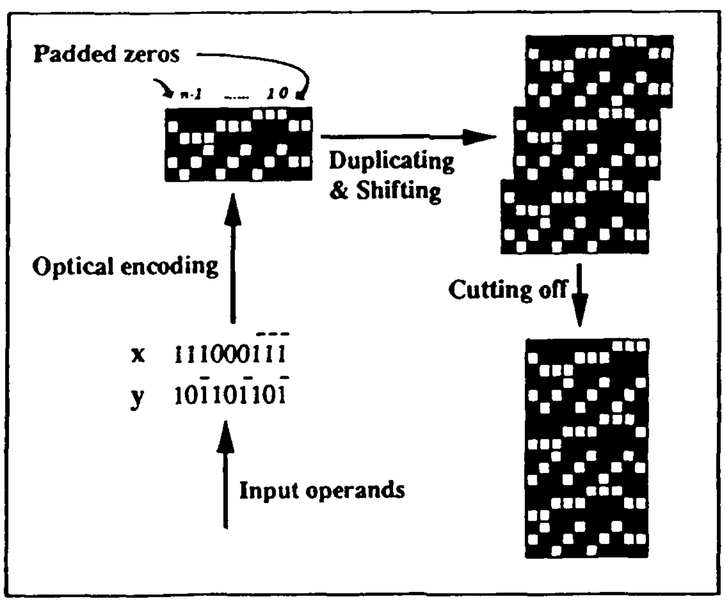


Fig. 6 Optical encoding procedure of input operands for MM implemention. Two n-bits operands are encoded into 18x(n+1) binary pattern.

Concepts for Fully Integrated FET SEED Circuits and their Applications for Optical Computing

A.J.Wachlowski*, F.A.P.Tooley, D.Rhein***

* Alcatel Austria Elin Research Centre, Ruthnergasse 1-7, A- 1210 Vienna, Austria
Tel.: +431/ 39 16 21, Fax.: +431/ 39 14 52.
** Physics Dept. Heriot-Watt University Riccarton, Edinburgh, UK EH14 4AS.

One approach to massively parallel and highly efficient information processing is to use systems containing multiple quantum well (MQW) modulators. The Symmetric Self-Electrooptic-Effect-Device (S-SEED) is one of the very well known examples for this technology [1]. The major mechanism of the MQW modulator, which is integrated in the intrinsic region of a PIN diode, is a clearly resolved exciton resonance at the absorption edge causing a strong nonlinearity in optical absorbtion. This exciton resonance shifts and broadens with electric field [2]. Although the S-SEED seems to be suitable for many applications in the field, for it is easy to use and capable for high switching speeds under some conditions, one main disadvantage appears to be the lack of logic integration. Even simple computing tasks require a large number of stages.

In our opinion it is neccessary to combine high speed semiconductor logic with integrated MQW modulators to make use of massively parallel free space interconnects to reach the processing performance of today's microprocessors. Doped channel HFETs or DMTs are suitable elements to realize the electronics part, as it is possible to integrate them together with multiple quantum well structures on one substrate [3]. This technology meets many demands for practical use: a high contrast of 6.5 and a small optical switching energy of about 1 pJ were reported [5]. The switching time was shown to be down to 2.8ns. [4] shows, how a 2x1 switching node has been realized in the FET SEED technology, where a ten times smaller optical switching energy than required for the S-SEED has been observed.

However, the FET SEED does not provide bistability as its predecessors did. So, we demonstrate a simple circuit to achieve electrooptical bistability. The circuit is shown in **Fig.1**. The circuit consists of two MQW diodes acting as input modulators and six HFETs forming a bistable flip-flop. We are still maintaining the use of dual rail beams to be compatible to S-SEED systems. In our simulation we used enhancement FETs for simulation purposes only, but it is easy to alter the circuit for the use of depletion types. In this case, the source potential has to be shifted, which is dependent on the pinch-off voltage of the actual type. The circuit, which

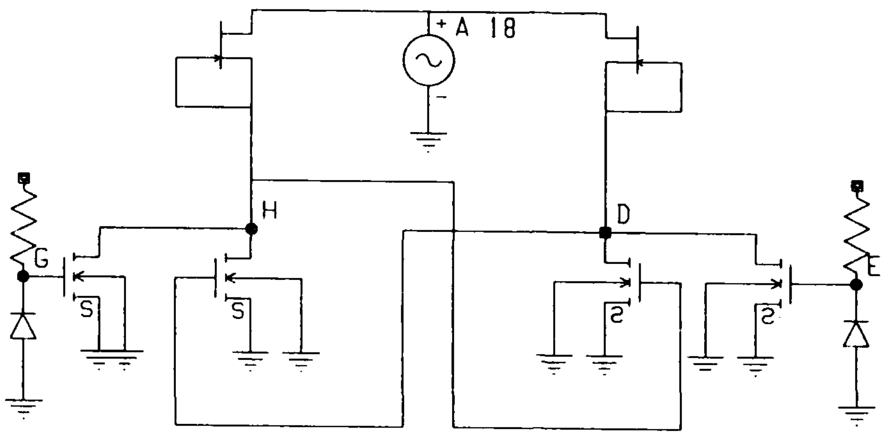


Figure 1: Bistable FET SEED circuit for input stages.
The MQW diodes are the optical input modulators, the nodes *H* and are the electrical outputs for internal use.

may be used as an input stage for several applications, provides two electrical complemented outputs. The width of the electrooptical hysteresis depends of the used FET parameters. Especially, the ratio width/length of the channel determines the switching levels of the stage. In this way, the electrooptical hysteresis can be optimized for optical computing applications by the variation of the FET channel parameters. The switching behaviour is shown in **Fig.2**. We used an analog workbench on a workstation to get the simulation results.

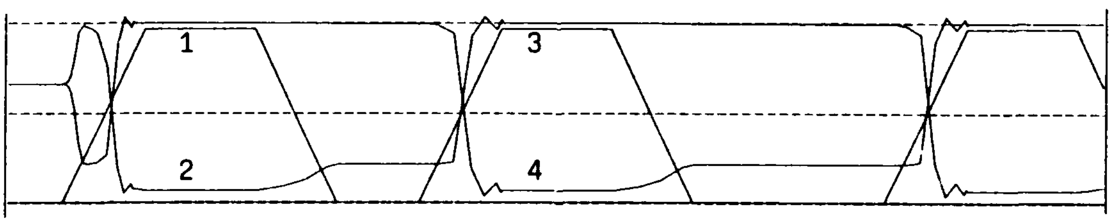


Figure 2: Switching properties of the bistable FET SEED circuit.
Referring to Figure 1, the input signals *1* and *3* are applied to the nodes *G* and *E*, while the outputs (nodes *H* and *D*) are represented by the signals *2* and *4*.

In the following two possible applications are discussed, which may be suitable for monolithic integration in large arrays for high parallelism. One example is a 2x2 switching node in extension to the 2x1 node described in [4], which consists of two input stages of the previously discussed type and four transmission stages. These transmission stages are actually simple NOR gates which deliver inverted output signals, but these are easily inverted in the output stage again. The node of the kind considered here is controlled by an electrical routing signal. However, other possibilities will be considered as well. The transmission stages allow two data paths between the input stages and the outputs of the circuit, "bypass" and "crossed". The circuit is illustrated in **Fig.3**. Twelf FET's are required for two bistable input stages and another twelf ones are required for the four transmission gates. Both the output stages are formed by six FET's. Over all, 30 FET's are used to create the complete circuit. It is also possible to cascade the 2x2 nodes to larger NxN nodes [6]. Another example for the use of FET SEEDs for optical computing purposes is the integrated, electro-optical full adder.The circuit consists of two input stages and a carry input and provides an

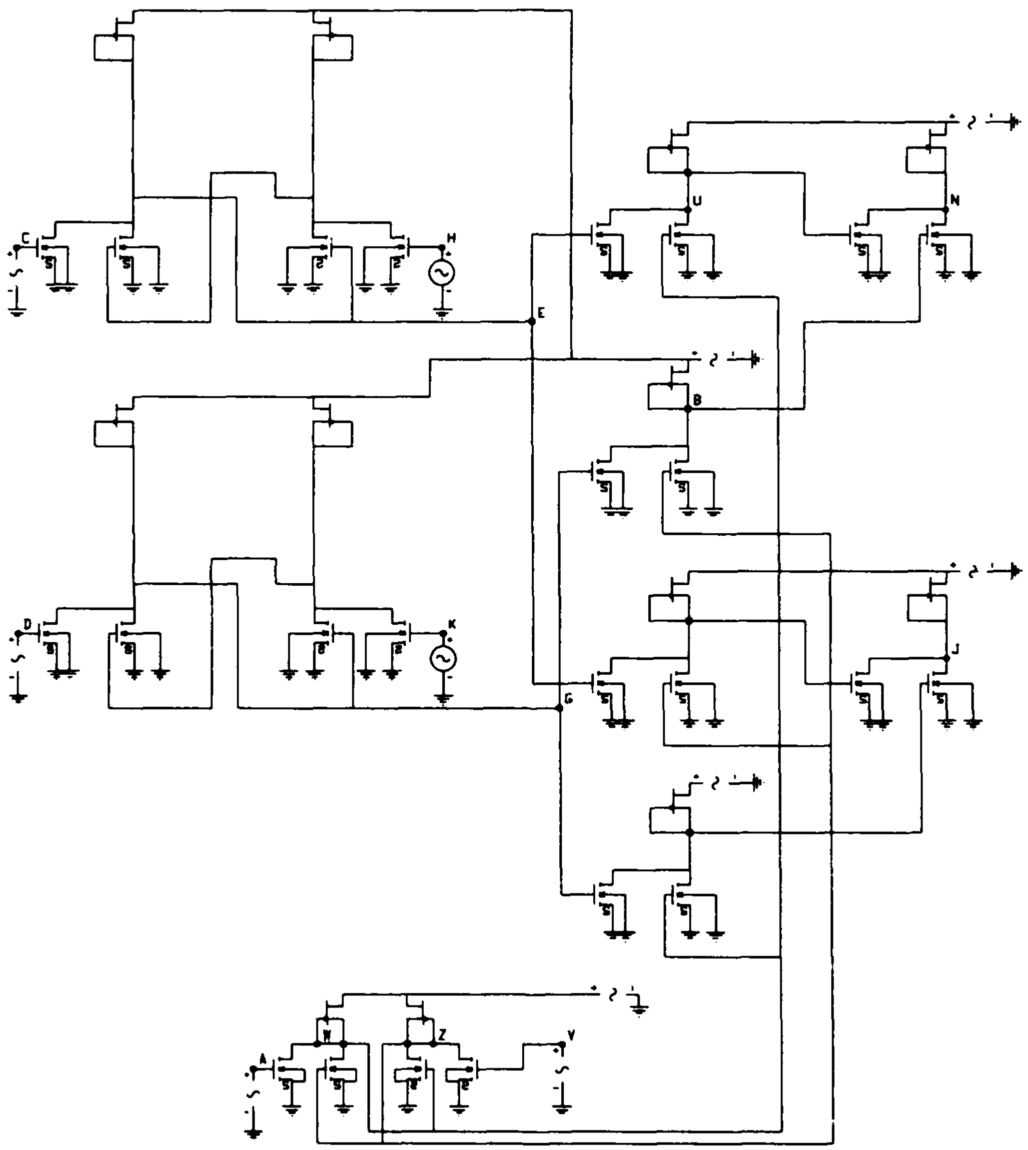


Figure 3: 2 x 2 switching node with FET SEED logic. It is built by two input stages, followed by four transmission gates and two outputs, represented by he nodes *N*, *J*.

optical output for the sum and the new carry. An unusual XOR gate is used to optimize the use of space. The simulation results will be discussed in detail. Considerations concerning timing and clocking are carried out to optimize performance. Over all, it is neccessary to consider the properties of asynchronous circuits. We will apply these considerations to both our examples.

References:

[1] D.A.B. Miller, D.S. Chemla, T.C. Damen, A.C. Gossard, W. Wiegmann, T.H. Wood, and C.A. Burrus, "Electric field dependence of optical absorption near the band gap of quantum-well structures", Physical Review Letters, Vol.53, No.22, 26.11.1984.

[2] A.L. Lentine, H.S. Hinton, D.A.B. Miller, J.E. Henry, J.E. Cunningham, and L.M.F. Chirovsky, "Symmetric self-Electrooptic effect device: optical set-reset latch, differential logic gate, and differential modulator/detector.", IEEE J. of Quantum Electronics, 25, 1928-1936, 1989.

[3] T.K. Woodward, L.M.F. Chirovsky, A.L. Lentine, L.A. D'Asaro, et al., "GaAs/AlGaAs FET-SEED smart pixels with several transistors", Summer Topical Meeting Digest, Smart Pixels 1992 Santa Barbara pp61-62.

[4] A.L. Lentine, T.J. Cloonan, and H.S. Hinton, " 4x4 arrays of FET-SEED embedded control 2x1 switching nodes", Summer Topical Meeting Digest, Smart Pixels 1992 Santa Barbara.

[5] T.K. Woodward, L.M.F. Chirovsky, A.L. Lentine, L.A. D'Asaro, M. Focht, et al., "Operation of a fully integrated GaAs-AlGaAs FET SEED: a basic optically addressed integrated circuit.", IEEE Phot.Techn.Lett., Vol.4, No.6, 614-617, 1992.

[6] M. Murdocca: " A digital design methodology for optical computing", MIT Press, Massachusetts Institute of Technology, Cambridge, Massachusetts 1990.

Strained multiple quantum well all-optical bistable microcavity operating at 980 nm

J. L. Oudar, D. Pellat, R. Kuszelewicz, and R. Azoulay

FRANCE TELECOM, Centre National d'Etudes des Télécommunications, Paris B,
Laboratoire de Bagneux,
196, Avenue Henri Ravera, BP 107, 92225 Bagneux Cedex, France
Tel: (33) 1 4231 7258
Fax: (33) 1 4253 4930

Quantum well vertical cavity structures are very attractive for the large scale implementation of parallelism in photonic systems, as they allow to fabricate 2-dimensional arrays of active functional devices such as surface emitting lasers, electro-optical modulators, and bistable optical switches. Nonlinear optics offers the possibility of all-optical operation, avoiding the cumbersome electrical wiring and addressing of large arrays.

Bistable étalons made of GaAs/AlGaAs Multiple Quantum Wells (MQW) exhibit very good nonlinear properties suitable for optical signal processing. Epitaxial methods allow to include the mirrors and the multiple quantum well nonlinear medium in a single crystal, resulting in highly compact microcavities, with typically 6 µm overall thickness. Optical bistability is observed in these structures at mW optical power, with a high contrast in the reflective mode[1].

We report on recent progress in

(1) reducing the power density threshold of all-optical bistable quantum well vertical microcavities. Significant improvements (by up to a factor 10) are achieved through an increase of the cavity finesse, together with a reduction of the device active layer thickness.

(2) implementing the concept of *resonant periodic nonlinearity* for an enhanced coupling of the nonlinear active quantum wells with the optical field of the cavity. This results in a still reduced number of quantum wells, which has allowed us to obtain all-optical bistability at 980 nm with strained InGaAs/GaAs quantum wells embedded in a high finesse AlAs/GaAs microcavity.

1. High finesse microcavities

Theoretical considerations taking into account the saturation of nonlinear refractive index in multiple quantum wells show that a significant reduction of bistability threshold intensity can be obtained through an increase of the cavity finesse, simultaneously with a decrease of the nonlinear layer thickness. The experimental demonstration of this feature is given through the comparison of several sample structures with different finesses and active layer thicknesses. Structure I, already described in Ref. 1, consists of 14 periods of $Ga_{0.9}Al_{0.1}As/AlAs$ quarter-wave thick layers acting as a back mirror, a 130-period MQW active medium consisting of 10 nm-thick GaAs wells and 10 nm-thick $Ga_{0.7}Al_{0.3}As$ barriers, and seven periods of the same $Ga_{0.9}Al_{0.1}As/AlAs$ alternate layers acting as a front mirror, grown on a semi-insulating GaAs substrate. Such a structure (sample #1896) was grown by Metal-Organic Vapor-Phase-Epitaxy (MOVPE). Its measured finesse of 36 was essentially equal to the theoretical value. It displayed a bistable behaviour with a threshold power of 2.8 mW and a high reflectivity contrast of 30:1.

Structure II consists of 23.5 periods of AlAs/$Ga_{0.9}Al_{0.1}As$ quarter-wave thick layers acting as a back mirror, a 18.5-period MQW active medium consisting of 10 nm-thick $Ga_{0.5}Al_{0.5}As$ barriers and 10 nm-thick GaAs wells, and 17 periods of the same AlAs/$Ga_{0.9}Al_{0.1}As$ alternate layers acting as a front mirror. The theoretical finesse of this structure lies in the interval 500-800, depending on the absorption coefficient of the nonlinear medium. Two samples of structure II (samples #FB17, and #GB23) were grown by Molecular Beam Epitaxy (MBE)[2]. On sample FB17, the measured finesse was 250, and the measured critical intensity for optical bistability was 5 $\mu W/\mu m^2$. On sample GB23, an additional quantum well in the active layer allowed to shift the microcavity resonance towards longer wavelengths, which resulted in a higher finesse of 380 at 864 nm. However the bistability threshold was not improved, but actually somewhat increased. Finally another sample (#2329) with a similar structure as structure II was grown by MOVPE. A better mirror quality resulted in a measured maximum finesse of 700 at still a longer wavelength of 871 nm. This resulted in an improved bistability threshold of 3.5 $\mu W/\mu m^2$. In addition, quite large hysteresis loops could be obtained on this sample, with a ratio of upper to lower threshold intensities up to 3:1.

Table 1 summarizes these results, and allows a visual comparison of the various sample characteristics. This shows that high finesse microcavities indeed allow a significant reduction of the bistability threshold. Together with an estimated carrier lifetime of 4 ns, the minimum threshold intensity corresponds to a power-lifetime product of 15 $fJ/\mu m^2$.

Sample		1896	FB17	GB23	2329
Mirror Reflectivities	R_b R_f	.976 .917	.998 .995	.998 .995	.998 .995
Number of GaAs QW		130	18	19	18
Finesse		36	250	380	700
Threshold ($\mu W/\mu m^2$)		40	5	6	3.5

Table 1: Comparison of bistability threshold in various GaAs quantum well samples

2. High finesse microcavity with strained InGaAs/GaAs nonlinear multiple quantum wells

Following the development of Vertical Cavity Surface Emitting Lasers (VCSEL), several InGaAs/GaAs nonlinear optical devices have been studied, such as an asymetric Fabry-Perot reflection modulator[3] or self-electro-optic effect devices[4], with operating wavelengths compatible with the recent InGaAs/GaAs VCSEL's. The substrate transparency at the operating wavelength offers new possibilities through the transmission of input and/or output beams, while avoiding any substrate etching and the hazardous sample manipulation that follows. For these reasons it was felt interesting to attempt observing all-optical bistability at 980 nm, by using such InGaAs/GaAs quantum wells as the nonlinear medium of a high finesse microcavity.

The structure design keeps the number of quantum wells as small as possible, in order to avoid any strain relaxation during the growth. This is obtained by using a high finesse microcavity as demonstrated above for GaAs wells, and a nonlinear medium in which the InGaAs quantum wells are located at the antinodes of the standing-wave optical field. This constitutes a resonant periodic nonlinearity, which enhances the coupling of the quantum wells with the intracavity field and allows to reduce the total InGaAs thickness.

The investigated sample (N° 2403) consists of 23.5 periods of AlAs/GaAs quarter-wave-thick layers as the back mirror. The nonlinear medium has a 5 half-wave optical thickness, with 4 groups of 3 $In_{0.15}Ga_{0.85}As$ quantum wells separated by GaAs barriers, with 10 nm/10 nm nominal thicknesses. GaAs spacers (88.3 nm thick) are grown between each MQW group in order to ensure the proper half-wave periodicity, as shown on Figure 1. Finally the front mirror consists of 17 periods of AlAs/GaAs alternate layers.

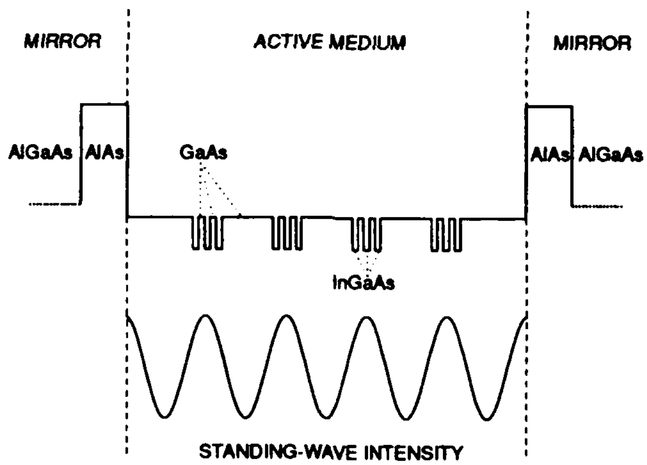


Figure 1: Structure of the active medium showing the InGaAs/GaAs quantum wells located at the antinodes of the standing-wave optical field.

The sample characterization by high resolution double X-ray diffraction and photoluminescence showed that the sample structure is very close to the designed values[5]. The reflectivity spectra exhibit an excitonic peak at 970 nm, and a cavity resonance in the range 978-985 nm, with a full width at half maximum of 0.5 nm at 980 nm. Taking into account the phase dispersion in the Bragg reflectors and in the nonlinear medium, we calculate a free spectral range of 80 nm, which implies a finesse of approximately 160.

The sample nonlinear behavior was studied using a Ti:sapphire laser operating in the long-wave range (920-1020nm). A 20 ns rise time acousto-optic modulator formed 1-µs-long triangular or rectangular pulses at the rate of 10 kHz. The beam was then focused on the étalon with a measured spot diameter of 24 μm (at half maximum) at normal incidence. When tuning the laser wavelength, we observed a clear bistable behavior at 983.4 nm, as displayed on Figure 2, with a reflective contrast of 7:1 and an intensity threshold of $10\mu W/\mu m^2$. This result is quite promising, in view of the potential improvement in materials quality and thickness controls in optimized samples. In the investigated sample the cavity loss is much larger than the value expected from the mirror composition. A better mirror quality is thus expected to improve the device performance and should also allow to obtain more detailed information on the nonlinear refractive index of these quantum wells.

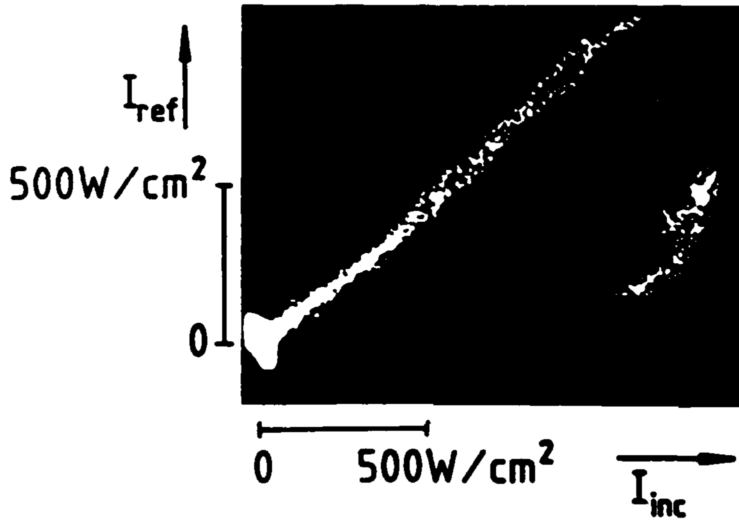


Figure 2: Reflected vs incident intensity showing the bistable behavior at 984.3nm.

In conclusion we have described significant improvements of bistability threshold in high finesse microcavities, and demonstrated a novel monolithic bistable device operating at about 980 nm, based on the dispersive nonlinearity of InGaAs/GaAs quantum wells. Its geometry and operating wavelength make it fully compatible with existing VCSEL arrays. Finally the substrate transparency at the operating wavelength opens new possibilities, e. g. for optical logic in the transmission mode, or for more efficient heat-sinking in the reflection mode, with incident and reflected beams propagated through the substrate.

1. B. G. Sfez, J. L. Oudar, J. C. Michel, R. Kuszelewicz, and R. Azoulay, Appl. Phys. Lett. **57**, 324 (1990).

2. J.-L. Oudar, R. Kuszelewicz, J.C. Michel and R. Planel, Electr. Lett. **28**, 714 (1992)

3. R.Jin, G. Khitrova, H.M. Gibbs, C. Lowry and N. Peyghambarian, Appl. Phys. Lett. **59**, 3216 (1991)

4. K. Fujiwara, Kenji Kawahima and K. Kobayashi, Appl.Phys.Lett, **57**, 2234 (1990)

5. D. Pellat, R. Azoulay, G. Leroux, L. Dugrand, Y. Rafflé, R. Kuszelewicz, and J. L. Oudar, submitted to Appl. Phys. Lett.

Smart Pixels: 1

OThA 8:30am-10:00am
Grand Ballroom West

Richard A. Linke, *Presider*
NEC Research Institute Inc.

Smart Pixel Optical Computing Architectures

A.A. Sawchuk and L. Cheng
Signal and Image Processing Institute, MC 2564, University of Southern California
Los Angeles, CA 90089-2564

S.R. Forrest and P.R. Prucnal
Advanced Technology Center for Photonics and Optoelectronic Materials (ATC/POEM)
Department of Electrical Engineering, Princeton University, Princeton, NJ 08544-5263

We describe five classes of smart pixel devices with increasing complexity and present an optoelectronic integrated circuit (OEIC) smart pixel array having programmable amplifier, inverter, logic element, bistable switch or latch functions. These devices are used in transmissive and reflective architectures to implement reconfigurable processors, multiplexed shuffle networks, cellular hypercube processors or sorting networks.

1.0 Introduction

Optics is advantageous for making intricate, long distance interconnections or transferring many signals on or off a chip. Electronics is better at power efficient signal amplification, processing and logic. Smart pixel (SP) devices have recently attracted great interest because they potentially combine the best characteristics of optics and electronics in a single device. SP devices are integrated optoelectronic arrays of light detectors, amplifying and processing electronics, light emitters and light modulators that are (ideally) combined on a single substrate. There are at least four major types of SP technologies being studied at present: a) multiple quantum well (MQW) self electro optic effect devices (SEEDs); b) optoelectronic integrated circuits (OEICs) with separate detectors, transistors, modulators, or optical sources in each pixel; c) vertically integrated detector, switch and emitter combinations; d) liquid crystal modulators integrated with VLSI drive and detection circuitry.

2.0 Complexity of Smart Pixel Device Arrays

We classify existing and proposed SP device arrays into five levels of complexity depending on their functionality. The first level are arrays of devices that detect light, perform an amplifying or thresholding function, and drive an output light source (LED or laser) or modulator at each pixel. These devices have a single input and output and perform a fixed function.

The second level are SP devices that can perform some signal routing or switching node functions. These devices can select from one or more inputs and route them to an optical output. Arrays of these devices can be controlled externally to build up switching networks with many inputs and outputs arranged in an array. One particularly useful type of device performs a bypass/exchange switching function. This operation takes a pair of inputs and routes them to the output in either a straight-through connection (bypass) or a reversed connection (exchange). This operation is a fundamental building block of multistage switching networks such as the shuffle (Omega), banyan, Clos/Benes, crossover, and many others [1].

The third level of complexity adds some degree of reconfigurability or controllability to the previous devices. The threshold level may be adjustable depending on an electrical or optical bias, and there may be additional signal inputs for choosing among several simple combinatorial logic functions such as AND, OR, NOR, etc.

The fourth level of complexity is a SP device that performs switching and routing functions as do the first three devices, but with some internal signal recognition and control. These smart pixels may require buffering of input signals and header, routing or destination recognition. These devices require a large number of transistors or logic devices at each node (perhaps several hundred), but a relatively smaller number of optical or electronic external input lines.

The fifth and highest level of complexity is a SP array consisting of individual processors capable of executing numerical or logical computations. Each of these smart pixels has individual optical and/or electronic input and output. There may also be local electronic connections on the chip to nearby processors for efficient short-distance data movement. These processors could perform parallel numerical signal processing functions such as the discrete Fourier transform, array beamforming computations, etc. They could also perform binary and multilevel image processing functions in a 2-D parallel array, as described by the digital optical cellular image processing (DOCIP) array concept [2], [3].

Note that SP arrays may function in either a transmissive mode (detectors and sources on opposite sides of the chip) or a reflective mode (detectors and sources on the same side of the chip). The device complexity, choice of fabrication technology and packaging with optical interconnection components will determine the most useful configuration.

2.1 A Multifunctional OEIC Smart Pixel. A particular OEIC SP device [4], [5] shown in Fig. 1 can be used in both transmissive and reflective systems. By optically changing the bias states on the control transistors (I_{c1}, I_{c2}, I_{c3}, I_{c4}, and I_{pb}), this circuit can be configured to provide several logic functions, provided that the optical inputs are simultaneously incident on transistor pairs (*IN*1, *IN*3) and (*IN*2, *IN*4). By providing positive optical feedback from the output laser to the control transistor, I_{pb}, the circuit can self-latch to hold an output value. Finally, pairs of such circuits can be employed to form a bypass/exchange switch.

Positive logic (AND, OR) is afforded by supplying base current to (i.e. enabling) *IN*3, *IN*4 and I_{c3}, I_{c4}, and negative logic (NAND, NOR, INVERT) is afforded by enabling *IN*1, *IN*2, I_{c1}, I_{c2} and I_{pb}. Here, a small signal current gain $\beta \sim 100$ - 500 is typical of heterojunction bipolar transistor (HBT) technology operated at ~1 GHz. The typical required base current values for the five control transistors are $1/\beta$ of the output laser threshold current I_{th} . To achieve optical reconfiguration of the circuit, the control transistors are optically sensitive elements such as phototransistors, photoconductors, or p-i-n photodiodes, depending on the performance and integration requirements of the system. One of the advantages of OEIC-based SP arrays in comparison to some of the other implementation technologies is that they have a large potential contrast (or on-off) ratio (> 5:1). This makes the OEIC devices better suited to applications requiring a larger fan-out and fan-in, which is particularly important for neural-based architectures, multiple input logic functions, and certain signal processing functions.

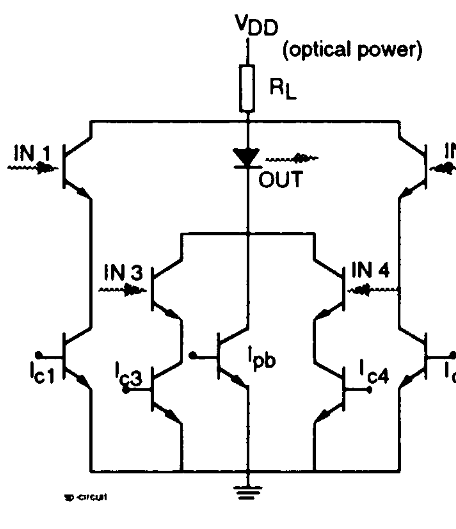


Figure 1. OEIC smart pixel circuit.

2.2 Transmissive Implementation. Figure 2 shows details of a general purpose optical computing system for transmissive versions of the OEIC SP arrays described previously. Here the latching functions L1 and L2 are shown separated from the SP arrays SP1 and SP2 for clarity, although they may actually be integrated along with the SP arrays. The arrays SP1 and SP2 can be reconfigured within a clock cycle to perform different Boolean logic functions or bypass/exchange spatial shifting or switching operations. The mirrors (M) shown route the array of signals through the system while the interconnection units perform fixed permutations on the array. Imaging elements needed to carry the data through the system and optical programming arrays for reconfiguration are omitted for clarity, but are discussed further in Ref. [5] and section 2.3. An input array enters L1 from the upper left, and proceeds through SP1 which is configured to perform a specific function on the first clock cycle. The result in stored in L2 while SP2 is configured for its operation. At the next clock cycle, SP2 processes its input and outputs the result to L1 while SP1 is reconfigured. The latches are used as temporary short-term memory during the SP computation time of ~ 1 ns or less. This time is short compared to the data transfer time of ~ 1 ps. The output array is obtained from the partially reflecting mirror shown.

2.3 Reflective Implementation. The concept shown in Fig. 2 with a *transmissive* SP device can be transformed to the *reflective* system shown in Fig. 3. Some advantages of the reflective system are that it may be more compact, easier to package and align, and requires SP arrays and interconnection holograms

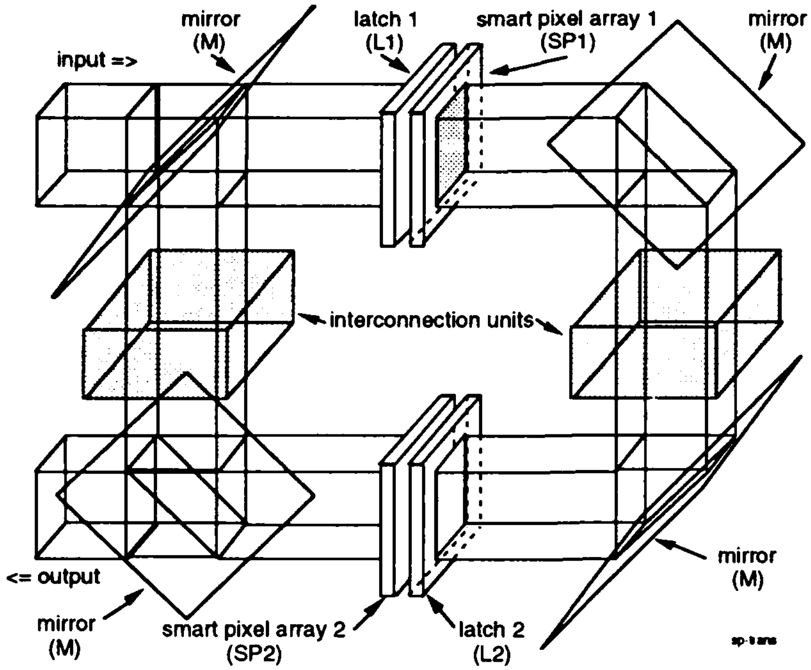


Figure 2. Transmissive smart pixel (SP) system.

that are easier to fabricate. Assume now that the latch (L) and smart pixel (SP) arrays are combined. The transformation begins by placing a vertical dividing plane through the planes of the L-SP arrays in Fig. 2 perpendicular to the forward and reverse optical paths. The components and paths to the left of the vertical dividing line are folded about the plane to the right side. Now place a horizontal plane through the interconnection units of Fig. 2, and fold the components below it upwards about the horizontal plane. The transmissive interconnection units of Fig. 2 become a reflective device, and the two SP arrays are now placed off-axis so that they mutually image to one another. Note that the interconnection unit is labeled as a hologram, but could be implemented using diffractive optics fabrication techniques. For imaging or simple shifting of one array with respect to another, the interconnection unit is a hologram with focal power that serves as a reflective imaging system. The single partially reflecting mirror that remains may be used to supply instructions, clock signals and provide input/output functions, or it may be eliminated if another program hologram placed next to or behind the reflective interconnection hologram provides these functions as shown in Fig. 3. An optical source behind the program hologram provides a clock and programming signals that pass through patterned transmitting regions of the interconnection hologram.

In the absence of a fully integrated SP array having detectors, electronic processing and sources on the same substrate, Fig. 3 shows how an array of detectors and processors forming a partial SP array can be electrically connected using wire bonds to a diode laser source array (such as a vertical cavity surface emitting laser (VCSEL) array). This arrangement may be useful for experimental feasibility studies until more fully integrated SP devices are available. Notice that simultaneous independent data transfers can occur from top (S1) to bottom (S2) SP arrays and vice-versa.

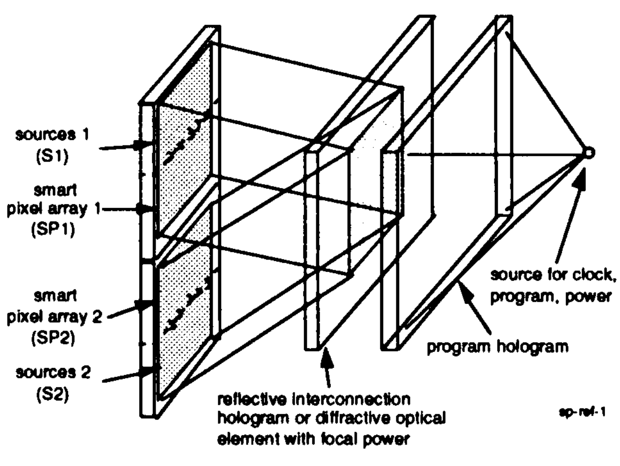


Figure 3. Reflective smart pixel system.

3.0 General Architectures

Here we describe four optical interconnection and computing systems that are well matched to the physical capabilities of the SP devices discussed in section 2.0 and can be implemented using the general SP array systems in Figs. 2 and 3. The systems are described in order of increasing complexity, primarily dependent on the complexity of the SP array, the fixed optical interconnection unit, input/output and programming functions required.

3.1 Reconfigurable Processor. The most straight forward version of the system in Fig. 2 and Fig. 3 requires a SP array of the logic elements described in section 2.1 along with simple imaging or space-invariant shifting interconnection optics. Here, the interconnections between pixels are one-to-one, with an entire array imaged to another array, perhaps with a spatial shift in x or y of a few pixels. Such a system could perform recursive logical operations on a binary array, including logical inversion, union of two arrays, or dilation (translation and superposition [2], [3]). Structures such as N-bit delay lines or shift registers could be implemented. Because of the 2-D array, several (up to N) parallel N-bit delay lines could be laid out on the SP chip.

3.2 Time or Space Multiplexed Shuffle. A perfect shuffle multistage network has in stages of 2-input/2-output bypass/exchange switches interlaced with fixed optical 2-D shuffle permutations [1], [6], [7]. The system of Fig. 3 can implement a time-multiplexed shuffle of synchronously processed data by transfers of data between the two halves of the SP array. Here the SP array is externally programmed to be a bypass/exchange array, and the optical interconnection is a fixed reflective perfect shuffle that provides one-to-one permutations from its input to its output. A SP array with a larger number of individual pixels could be used to space-multiplex some or all of the stages [8], thus reducing the required number of clock cycles (data transfers).

3.3 Cellular Hypercube (CH). The CH is an enhanced binary logic array that adds direct holographic interconnections to pixels at distances 2, 4, 8, ..., 2^k ($2^k \leq N/2$) (a one-to-many fan-out) in order to reduce the number of computation clock cycles [2], [3]. For an array of size $N \times N$, the fan-out from the source at each pixel is $O(\log N)$ rather than one as in system 3.1 and 3.2. Here, the same basic logic SP as in Fig. 1 is used.

3.4 Sorting Network. The most advanced application of these SP systems is in sorting or self-routing interconnection networks. The smart pixels must perform logical operations such as comparisons, header reading and adaptive routing of data passing through the system. The basic interconnection between stages is a one-to-one interconnection such as a shuffle or other multistage network connection.

4.0 Other Issues, Conclusions and Acknowledgments

Additional optical sources and photovoltaic detectors may be added to the SP systems in Figs. 2 and 3 to provide optical powering of the SP arrays. Eliminating electrical power supply connections can significantly reduce inter-pixel crosstalk and interference [8]. With present technology, the packing density of MQW SEED SP arrays is higher than that for OEIC SP arrays, but the fan-in, fan-out, and data signal bandwidth potential of OEICs appears much greater. The best choice of SP will depend on the ease of packaging and physical integration with wavelength-compatible diffractive optics for the fixed interconnection units.

This work was supported by NSF Grant No. ECS-9015797 and by the Air Force Office of Scientific Research under Grant No. F49620-92-J-0432.

5.0 References

1. L. Cheng and A.A. Sawchuk, "3-D Omega Networks for Optical Implementation," *Applied Optics*, vol. 31, pp. 5468-5479, (1992).
2. K.-S. Huang, B.K. Jenkins and A.A. Sawchuk, "Binary Image Algebra and Optical Cellular Logic Design," *Computer Vision, Graphics and Image Processing*, vol. 45, pp. 295-345, (1989).
3. K.-S. Huang, B.K. Jenkins and A.A. Sawchuk, "Image Algebra Representation of Parallel Optical Binary Arithmetic," *Applied Optics*, vol. 28, pp. 1263-1278, (15 March 1989)
4. J.J. Brown, J.T. Gardner, and S.R. Forrest, "An Integrated, Optically Powered Optoelectronic 'Smart' Logic Pixel for Interconnection and Computing Applications," to appear in *IEEE Journal of Quantum Electronics*, (1993).
5. S. Yu, S.R. Forrest, and P.R. Prucnal, "An Optoelectronic Processor/Interconnection Employing Smart Logic and Routing Pixel Arrays," submitted to *IEEE Photonics Technology Letters*, (1992).
6. L. Cheng and A.A. Sawchuk, "A One-Copy Algorithm for 2-D Shuffles for Optical Omega Networks," *Journal of Parallel and Distributed Computing*, vol. 16, pp. 54-66, (1992).
7. J.-M. Wang, L. Cheng and A.A. Sawchuk, "Optical 2-D Perfect Shuffles Based on a One-Copy Algorithm," *Applied Optics*, vol. 31, pp. 5464-5467, (1992).
8. M. Govindarajan, S.R. Forrest, L. Cheng and A.A. Sawchuk, "Optically Powered 2x2 Optoelectronic Switch with Polarization Routing," *IEEE Photonics Technology Letters*, vol. 3, pp. 669-672, (1991).
9. L. Cheng, J.-M. Wang and A.A. Sawchuk, "Design of Space-Multiplexed Three-Dimensional Omega Networks and their Optical Implementation," submitted to *Applied Optics*, (1992).

Field Effect Transistor — Self Electrooptic Effect Device (FET-SEED) Circuits for Optoelectronic Data Processing Systems

L. M. F. Chirovsky, L. A. D'Asaro, E. J. Laskowski, and S. S. Pei
AT&T Bell Laboratories, Murray Hill, N.J. 07974

M. T. Asom, M. W. Focht, J. M. Freund, G. D. Guth, R. E. Leibenguth, G. Livescu, R. A. Morgan, T. Mullaly and L. E. Smith
AT&T Bell Laboratories, Breinigsville, PA 18031

A. L. Lentine
AT&T Bell Laboratories, Naperville, IL 60566

G. D. Boyd and T. K. Woodward
AT&T Bell Laboratories, Holmdel, NJ 07733

Electronic chips have tremendous data processing capabilities but they also have a communications bottleneck. The amount of data that can be processed on electronic chips far exceeds the amount of data that can be brought in and out of these chips. This imbalance is due to the fact that high speed electronic data input/output (I/O) ports consume far more chip area and power than equivalent data processing circuits do. The imbalance is tolerable in many applications where a given data set undergoes many processing steps before leaving the chip. Then data can be pipelined through the multiplicity of steps or repeatedly routed through an on-chip cache memory. Such procedures are not efficient however, in applications where very large amounts of data need to undergo only a limited number of processing steps, and then input/output limitations belie the processing power of the electronic chips.

Recently, optical data processing prototype systems [1,2] have demonstrated that the large space-bandwidth product of free-space optics offers a means of vastly increasing the number of input/output ports on a chip. These systems utilized large arrays of Symmetric Self Electrooptic Effect Devices (S-SEED's) as their active components [3-5]. The devices acted simultaneously as data processing nodes and as the optical input and the optical output ports. Although the S-SEED's were surprisingly functionally versatile for elementary devices and exhibited the many device features that made the operation of the prototype systems possible [6], the S-SEED's had two shortcomings which stymied further progress. Their switching speeds in optically cascaded operation fell far short of their intrinsic speeds [7,8] and their data processing ability was at most that of a single logic gate. Two developments, however, pointed the direction towards solutions to the two problems. Both developments were possible because although S-SEED's act as optical devices, they are fundamentally optically addressed electronic circuits, consisting of two, reverse-biased in series, Quantum Well diodes (QW-diodes) which can be both detectors and modulators [9]. In the first development, a series of experiments showed that S-SEED's can be electrically driven and optically read at Gigahertz rates [5,10]. So the speed problem was not a fundamental one, but only a question of sensitivity to optical signals. The second development was a family of devices (actually circuits), called Logic-SEED's or L-SEED's [11] which achieved more complex functionality through higher circuit complexity. L-SEED's consist of many electrically interconnected QW-diodes. To operate, the L-SEED's need many optical beams (at least one per QW-diode), so they have a beam complexity proportional to the circuit complexity; not a desirable situation if ever higher functional complexity is to be explored. So the need for higher sensitivity and the need to disengage circuit complexity from beam complexity, both begged for an introduction of electronic components, such as transistors into SEED circuits. That QW-diodes could interact beneficially with transistors was obvious, but how to monolithically integrate, or cofabricate, QW-diodes with transistors, in a manufacturable and reliable fashion, was not.

We chose to develop a cofabrication sequence for QW-diodes and Heterostructure Field Effect Transistors (HFET's) since feasibility was already demonstrated [12] and because FET circuits are more likely to aid high sensitivity than, for instance, Heterostructure Bipolar Transistor (HBT) circuits do. (Efficient cofabrication of QW-diodes and HBT's is also possible [13].) Reference [14] gives a detailed description of a FET-SEED batch fabrication procedure; one that is manufacturable and allows versatile circuit design. The process requires only one more photolithographic step than a previously developed SEED batch fabrication sequence [15] and still involves a single MBE growth only slightly modified from those for just QW-diodes. Exemplary FET characteristics are amply described in reference [14]. For FET-SEED's we also modified the Multiple Quantum Well Stack (MQWS) of the QW-diodes. This change optimized

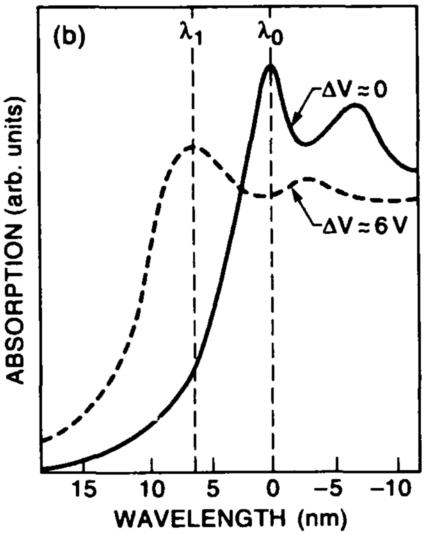


Figure 1. An illustration of the Quantum Confined Stark Effect, including a designation of two wavelengths, λ_0 and λ_1, useful for SEED operations.

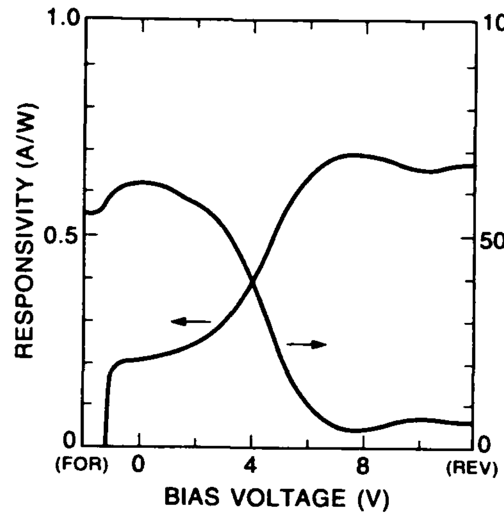


Figure 2. Plots of the responsivity and the reflectivity vs. applied voltage of a QW-diode, for light at λ_1 at an intensity of 10 kW/cm^2.

the QW-diodes for operation with light beams at the wavelength, λ_1, identified in Figure 1. The figure is an illustration of the Quantum Confined Stark Effect, the phenomenon which gives QW-diodes their novel properties. Previously, SEED's were operated at wavelengths near λ_0 (Fig. 1) to obtain bistability. But FET-SEED circuits do not need bistability and so can take advantage of the deeper modulations at higher optical beam intensities that are possible at λ_1. Figure 2 shows typical reflectivity and responsivity versus voltage characteristics obtained at λ_1 with a 10 kW/cm^2 beam intensity, for a Fabry-Perot resonance enhanced QW-diode. Even without the Fabry-Perot enhancement similar modulations, with the minimum reflectivity rising to 11%, can be achieved.

With the FET-SEED integration, all the components needed for optoelectronic circuits are available. But careful circuit design is still necessary to properly exploit the opportunities offered. Chips which utilize the high bit-rate massive connectivity of free-space optics and have a high level of functional complexity, must have receiver circuits which convert optical signals to electrical logic level signals, a family of electronic or optoelectronic logic circuits and transmitter circuits which convert logic level electrical signals into optical signals. For cascadability between chips, the optical outputs of the transmitter circuits must be able to drive the receiver circuits. Both receivers and transmitters must be relatively high speed, because high data throughput requires a large number of I/O ports which are also fast. Both receivers and transmitters must be relatively small if the I/O ports are not to consume a disproportionate amount of chip area and power. Thus our first circuit design task was to develop the proper receiver and transmitter circuits.

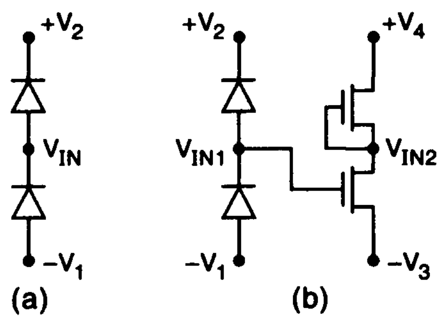


Figure 3. Schematic diagrams of FET-SEED receiver circuits.

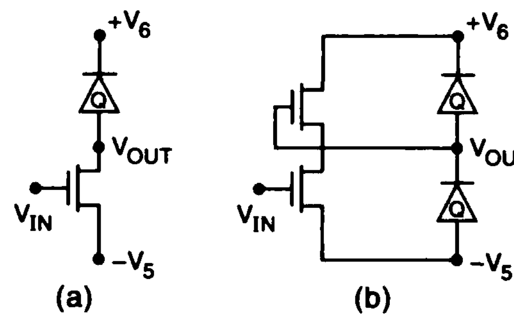


Figure 4. Schematic diagrams of FET-SEED transmitter circuits.

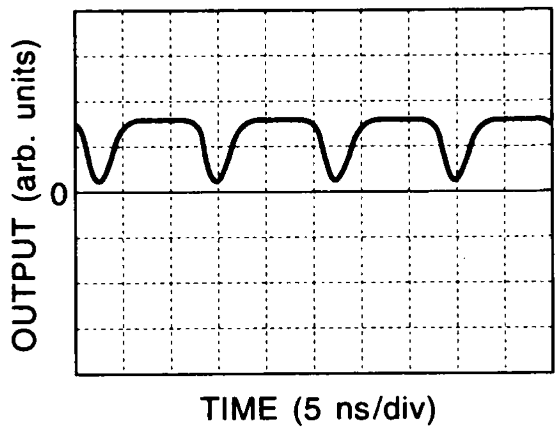


Figure 5. Optical output of a FET-SEED circuit consisting of a Fig. 3a receiver and a Fig. 4a transmitter. Test conditions were a CW read beam and every 12 ns, mode-locked laser pulse streams directed to the upper (set) diode and lower (reset) diode. Reset 2 ns after set.

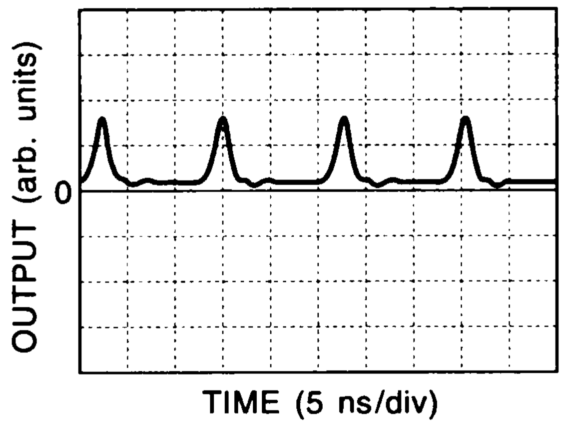


Figure 6. Optical output of a FET-SEED circuit consisting of a Fig. 3a receiver and a Fig. 4a transmitter. Test conditions were a CW read beam and every 12 ns, mode-locked laser pulse streams directed to the upper (set) diode and lower reset diode. Set 2 ns after reset.

The development strategy was — design variations of receivers and transmitters simultaneously, then layout various on-chip combinations, with receivers directly electrically driving transmitters, and finally test the combinations optically, using actual or simulated optical outputs of transmitters to drive the receivers and observing the transmitter optical outputs. A description of some examples follows.

The simplest receiver is a photodiode (a QW-diode in our case) reverse-biased in series with a load. The highest sensitivity such receiver occurs when the load is another photo-diode as shown in Figure 3a. This receiver can be operated with differential optical signals (with sensitivity increasing with increasing contrast) or with single-ended signals on one diode and time-sequentially applied reset pulses on the other. Some form of clamping of the voltage swing on V_{IN} to no more than a logic level swing must be added. When the receiver drives a gate of a FET, that gate can automatically be the clamp. A somewhat larger but more sensitive receiver is shown in Figure 3b. A full logic level swing on V_{IN2} can be obtained with a much smaller voltage swing on V_{IN1}. Some form of clamping on V_{IN1} must be provided also. Sensitivity can be measured in terms of the inverse of the number of photons needed to develop a logic swing at the receiver output. The switching energy (the number of photons times the energy per photon) is proportional to the voltage swing the optical signal must provide, so sensitivity increases as the required input voltage swing decreases.

The simplest and least power dissipating transmitter design is shown in Figure 4a. This transmitter gives only single-ended outputs and so must have high contrast to be cascadable with the receivers of Figure 3. When operated with pulsed read beams this transmitter draws current only when the read beam is on and uses read beam generated photo-carriers to reset itself. The transmitter in Figure 4b. is much easier to use and can give differential outputs, but simply dissipates the absorbed optical power and dissipates more electrical power because of the contention between the FET's.

A circuit consisting of a Figure 3a receiver and a Figure 4a transmitter was tested with mode-locked laser pulse streams directed at the input diodes (upper diode is set, lower is reset) and with a CW read beam on the output diode. Figure 5 shows the reflected output when the reset pulses trailed the set pulses by 2 ns (pulses were ~12 ns apart in both streams). Figure 6 shows the reflected output when the set pulses trailed by 2 ns. The two experiments together demonstrated that operation up to 250 Mbit/sec is possible. The energy in the set pulses was 80 fJ. The circuit may actually be faster, the test equipment limited the speed in those experiments. Another circuit, consisting of a Fig. 3b receiver (with diode clamps added) and a Figure 4b transmitter was tested in a similar fashion [16]. The results, shown in Figure 7, were a 200 Megahertz speed of operation with the average switching energy reduced to ~20 fJ.

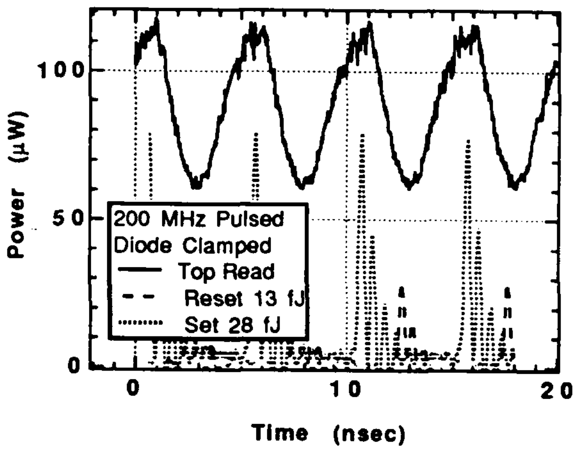
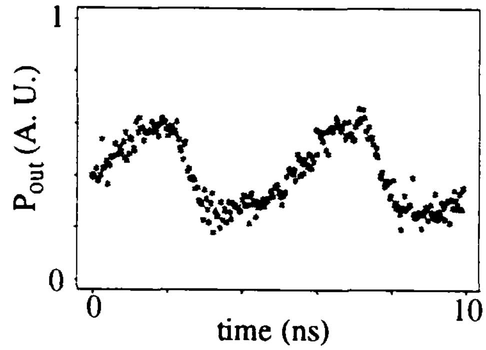


Figure 7. Optical output of a FET-SEED circuit consisting of a Fig. 3b receiver and a Fig. 4b transmitter. Test conditions were a CW read beam and every 5 ns, pulse streams (subnanosecond pulses) directed at the set and reset diodes, 2.5 ns apart.



13 μW/input beam
100 μW/output beam
input optical energy per beam = 60 fJ

Figure 8. Optical output of a complex FET-SEED data routing node. Test conditions were a CW read beam and two 200 MHz, 50% duty-cycle input beams, directed out of phase at the set and reset diodes of one of the node receivers.

The second generation of FET-SEED circuits already included some data processing nodes, specifically data routing nodes [17], processed in 4×4 arrays. These nodes included logic circuits, based on the "Buffered FET Logic" (BFL) family, placed between the receivers and transmitters. We have fully functional arrays, each with a total of 400 FET's and 304 QW-diodes (25 FET's and 19 QW-diodes per node). Figure 8 shows a 200 MHz output of one such node, receiving 60 fJ input pulses (at 50% duty cycle).

Despite the above successes, the FET-SEED technology is still facing many unanswered questions. The most critical one is to what complexity will yield allow the circuits to go. Perhaps the most exciting and stimulating question is what are the optical computing applications for these optoelectronic circuits. Remember that with FET-SEED circuits, optics is offering not only massive connectivity for unprecedented data throughput, but also automatic signal regeneration, signal retiming and clock distribution. The electronics bring arbitrary functional complexity.

References

[1] F. B. McCormick et al., Appl. Optics, V. 31, 5431-5446 (1992).
[2] M. E. Prise et al., Appl. Optics, V. 30, 2287-2296 (1991).
[3] A. L. Lentine et al., IEEE J. Quantum Electron., QE 25, 1928-1936 (1989).
[4] A. L. Lentine et al., IEEE Photon. Tech. Lett., V 2, #1, 51-53 (1990).
[5] L. M. F. Chirovsky et al., OSA Proceedings on Photonic Switching, H.S. Hinton and J. W. Goodman, eds. (OSA, Washington, D.C., 1991) V. 8, p. 48.
[6] L. M. F. Chirovsky et al. SPIE Proceedings, V. 1562, Devices for Optical Processing (1991) p. 228.
[7] A.L.Lentine et al., Optics Lett., V. 16, #1, 36-38 (1991).
[8] G. D. Boyd et al., Appl. Phys. Lett., V. 57, #18, 1843-1845 (1990).
[9] D. A. B. Miller et al., Appl. Phys. Lett., V. 45, p. 13 (1984).
[10] A. L. Lentine et al., IEEE Photon. Tech. Lett., V. 2, 477-480 (1990).
[11] A. L. Lentine et al., IEEE J. Quantum Electron., V. 28, 1539-1553 (1992).
[12] D. A. B. Miller et al., IEEE Photon. Tech. Lett., V. 1, 62-64 (1989).
[13] K. W. Goossen et al., IEEE Photon. Tech. Lett., V. 4, 393-395 (1992).
[14] L. A. D'Asaro et al., to be published in 1992 special issue on Smart Pixels of IEEE J. Quant. Electron.
[15] L. A. D'Asaro et al., *Advanced Processing and Characterization Technologies*, P. J. Holloway, ed., Am. Vac. Soc. #18, 1843-1845 (1990).
[16] T. K. Woodward et al., 1992 IEDM Tech. Digest, to appear in Dec., 1992.
[17] A. L. Lentine et al., Postdeadline Paper #1, IEEE Summer Topical Meeting on Smart Pixels, 1992.

Cascaded Operation of Two 128-Pixel FET-SEED Smart Pixel Arrays

F. B. McCormick, A. L. Lentine, and L. M. F. Chirovsky*
AT&T Bell Laboratories, Naperville, IL 60566
**AT&T Bell Laboratories, Murray Hill, New Jersey, 07974*

1. Introduction The current trend in device design for free-space optically interconnected systems is towards 2-D arrays of smart pixels, in which each pixel incorporates electronic processing together with optical I/O. The optimal amount of electronic functionality of each smart pixel is a topic of current debate, however, amplification of the detector output signals is common to most proposals. This amplification compensates for the typical optical systems losses of 3 to 15 dB which would otherwise limit the systems to low data rates. Progress has been made in fabricating smart pixel arrays incorporating GaAs FETs and MQW PIN detectors and modulators (FET-SEEDs)[1] and this paper presents an experiment which cascades two 128 pixel arrays of FET-SEED smart pixels.

2. POET FET-SEEDs The FET-SEEDs used in this experiment form pixels which consist of 2 MQW PIN photodiodes, a single-stage FET amplifier, and single-ended MQW output modulator (as shown in Fig. 1). The upper photodiode on the left is used for the signal pulse, the lower photodiode is used for the reset pulse. The output modulator is addressed with an output (power) pulse much larger than the input, thus reading out the state of the circuit and providing optical power gain. These pixels were originally designed to be operated with mode-locked pulses, thus their acronym: Pulsed Opto-Electronic Toggle, or POET. The zero-field excitonic absorption peak of these MQW PIN diodes is at ≈ 846 nm, and they were designed to operated at the so called λ_1 wavelength of 850 nm, where increasing field causes increasing absorption, hence the differential inputs are not bistable. The FET Shottky diode effectively "clamps" the reset diode to less than 1 volt, thus its absorption (≈ 45%) is always less than that of the signal diode (≈ 75%) at 850 nm.

3. Experiment A photograph and schematic of the experimental setup are shown in Fig. 2. The signal and reset signals for $POET_1$ are generated by AlGaAs laser diodes driven by a pulse generator. The reset laser was operated at 846 nm to maximize its absorption by the reset diodes, while the rest of the lasers were operated at 850 nm. These beams are combined with a 50:50 beam splitter, BS_1, after which a binary phase grating, BPG_1, splits this pair of beams into 128 pairs. These beams enter an optical isolator, ISO_1, are reflected by a polarizing beam splitter (PBS), pass through a $\lambda/4$ retarder with its fast axis at 45 degrees to the PBS plane of incidence, reflect from a 50% mirror, pass back through the PBS and another (similarly oriented) $\lambda/4$ retarder, and are imaged onto the input PIN diodes of $POET_1$.

The $POET_1$ output modulators are read using a Ti:Sapphire laser split into 128 beams by another BPG. These power beams pass through ISO_1, are imaged onto $POET_1$, and the reflected beams become the signals for $POET_2$. The path for these signals through ISO_2 to $POET_2$ is similar to the input signal path described above. The reset and power beams for $POET_2$ come from CW AlGaAs laser diodes and, after being split 128 ways by BPGs, they are combined together at BS_2. Both beam sets propagate through ISO_2 and are imaged onto $POET_2$ along a path similar to the path of the $POET_1$ power beams.

Much of the optics and mechanics used in this experiment are identical to those used in our last optical switching demonstrator,[2] and the alignment proceeds similarly. The semiconductor lasers and collimating/circularization optics were mounted on separate sub-plates, and then mounted onto the main custom steel mounting plate, which was milled with 7 shallow slots. Most of the other components were mounted in 25mm cylindrical steel cells held in a steel slot-plate with ceramic magnets. The FET-SEEDs were centered on cylindrical steel slugs, and the 3 spot arrays for $POET_1$ were aligned by rotating the BPGs and the pairs of 7 minute wedges (Risley prisms). The reflected $POET_1$ output beams were similarly centered on $POET_2$, and the $POET_2$ cell rotated to match $POET_1$. The $POET_2$ power and reset beams were then aligned to $POET_2$ by rotation of their BPGs and wedges. Alignment of the power beams was optimized by maximizing the average array photocurrent generated with only the power beams present. The signal and reset beams' alignment was similarly optimized by maximizing or minimizing the array photocurrent with both the power beams and either the signal or reset beams present.

The signal and reset beams for $POET_1$ were modulated with a 50% duty cycle and the $POET_1$ power was CW. The $POET_1$ array was operated at a 50 MHz with 100% functionality, although the output pulse shape and contrast was degraded from its DC performance. The input was a 10101.... pattern, and the resulting optical outputs from one quadrant of the array (32 pixels) are shown in Fig. 3(a). Figure 3(b) shows an overlayed plot of all 128 $POET_1$ outputs. These outputs were used to drive the signal inputs of $POET_2$, and the cascaded system was operated at ≈75% functionality at 14 MHz. The outputs from one quadrant of $POET_2$ shown in Fig. 4(a), and the overlayed plot of all 128 $POET_2$ outputs (Fig. 4(b)) demonstrate the broad range of observed performance. At 2.5 MHz, this range of performance was still evident, with 10 devices not switching, and 23 others switching with asymmetric duty cycles or very low contrast. The power emitted by each laser, as well as the optical power/pixel of signal, reset, and power beams is summarized in Table 1.

4. Conclusions The non-uniform performance of the cascaded system may be attributed to the effects of device non-uniformity and optical system non-uniformity. Figure 5 shows the near-DC switching response of every other device (1,3,5...) in every other row (1,3,5..) of the array, when sequentially operated with CW reset and power beams and a ramped-power signal beam. Only one of the 32 devices tested was non-functional, and some of the small non-uniformity present is due to misalignment errors from the automated positioning stages.

Optical signal non-uniformity is contributed by the BPGs, beam clipping at the detectors and modulators due to misalignment, as well as other smaller effects.[3] In similar systems, [2] optical effects typically introduce about 10% non-uniformity, and alignment effects another 10-20%. This 20-30% variation in signal powers can have a dramatic effect on the device switching dynamics across the array. Assuming a 20% variation about the nominal power, the LOW signal power/pixel for $POET_2$ may vary from 9.6 μW to 6.4 μW. Due to the decreased absorption of the $reset_2$ beams, the effective nominal power/pixel is 18(45/75)=10.8 μW. These beams may vary ± 10%, or from 11.9 μW to 9.7 μW. This variation is significant, because the FET gate-source voltage depends on the difference between the photocurrents generated by each diode of the differential detector, and hence the absorbed differential optical power, which may vary from $\Delta_{LOW-min} = 0.1 \mu W$ to $\Delta_{LOW-max} = 5.5 \mu W$. This represents a factor of 55 in resultant switching speed variation across the array! Additional array operational effects such as simultaneous switching noise, heating and scattered light effects may be present and will require further investigation. This work represents but a first step, and future experiments with pulsed inputs, higher contrast ratio devices, and differential inputs should yield improved system data rates, even in the presence of these non-uniformities.

We would like thank the authors of reference 1 for making the devices, and Rick Morrison and Sonya Walker for designing and fabricating the BPGs used in this work.

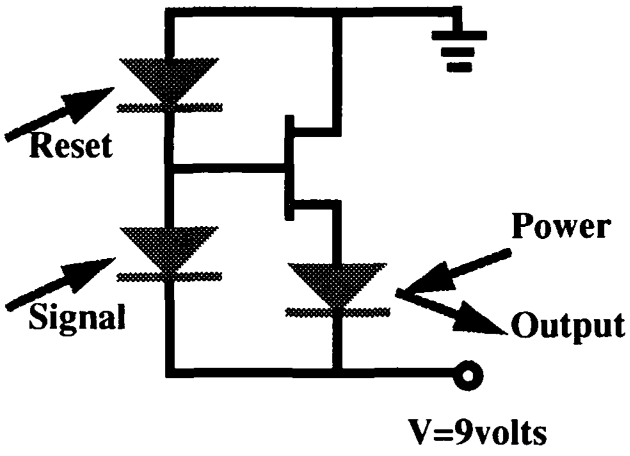
REFERENCES

1. L. A. D'Asaro, L. M. F. Chirovsky, E. J. Laskowski, S. S. Pei T. K. Woodward, A. L. Lentine, R. E. Leibenguth, M. W. Focht, J. M. Freund, G. Guth, and L. E. Smith, "Batch fabrication and testing of GaAs-AlGaAs field effect transistor self electro-optic effect device (FET-SEED) smart pixel arrays," IEEE J. Quantum Electron. (submitted for publication)

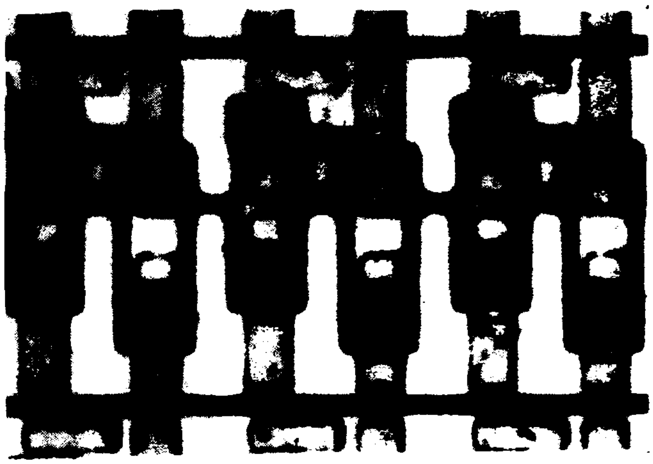
2. F. B. McCormick, T. J. Cloonan, F. A. P. Tooley, A. L. Lentine, J. M. Sasian, J. L. Brubaker, R. L. Morrison, S. L. Walker, R. J. Crisci, R. A. Novotny, S. J. Hinterlong, H. S. Hinton, and E. Kerbis, "A Six-Stage Digital Free-Space Optical Switching Network Using S-SEEDs," accepted for publication in *Applied Optics* , Oct. (1992).

3. F. B. McCormick, F. A. P. Tooley, J. L. Brubaker, J. M. Sasian, T. J. Cloonan, A. L. Lentine, S. J. Hinterlong, and M. J. Herron, "Optomechanics of a free-space switch: the system", *Optomechanics and Dimensional Stability*, Proceedings of SPIE, **1533**, 97-114, (1991).

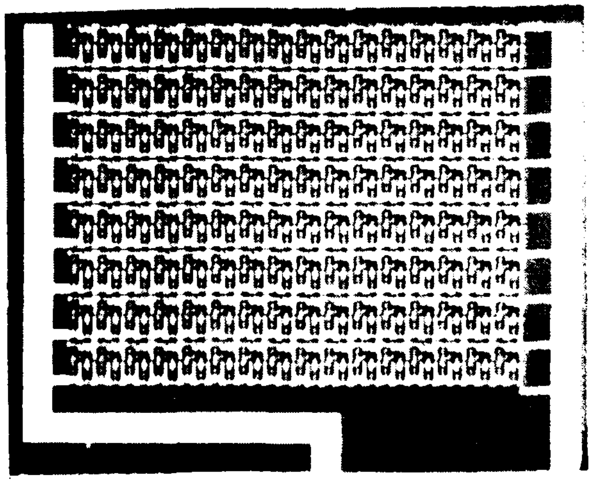
TABLE 1.	$reset_1$	$signal_1$	$power_1$	$reset_2$	$signal_2$	$power_2$
Total power at laser-pen (or $POET_1$ output) (mW)	21	31	42	22.5	HI=5.5 LOW=2.56	70
Power per pixel at each pixel (μW)	20	29	78	18	HI=17 LOW=8	99



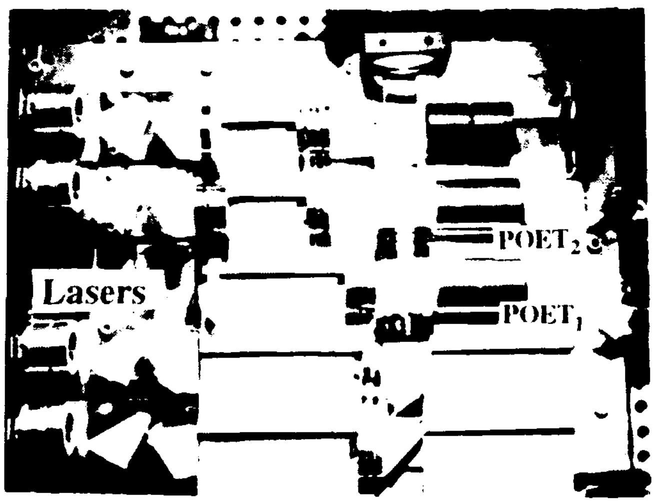
1. POET FET-SEED electrical schematic



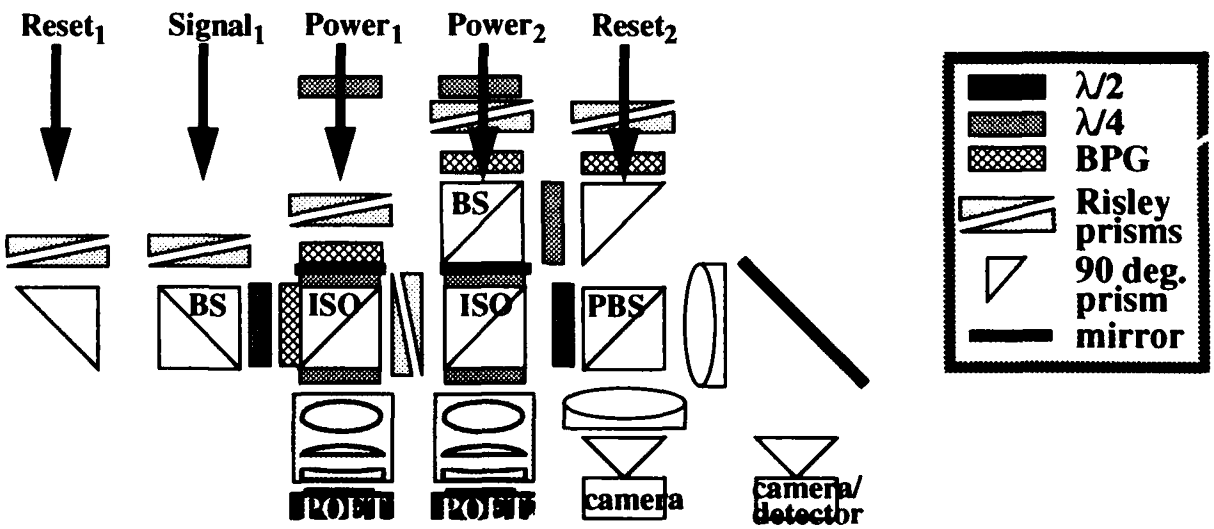
1(b).Three POET smart pixel cells



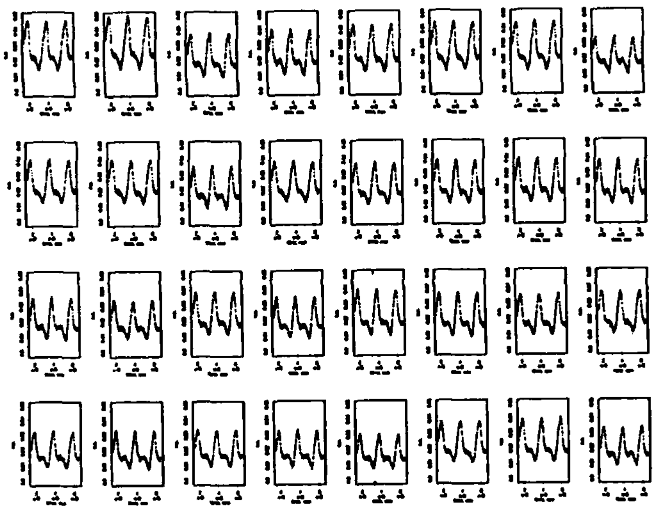
1(c). 16x8 POET array



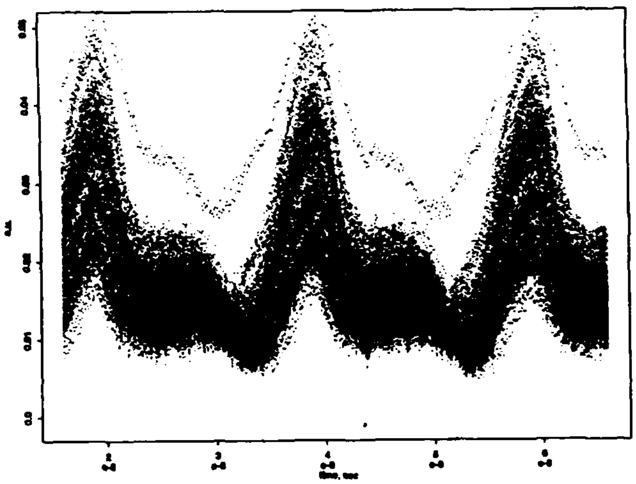
2(a). Experimental hardware: setup



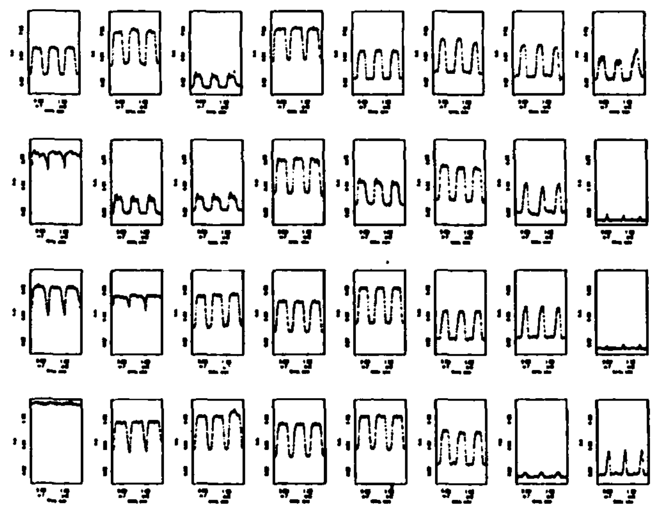
2(b). Experimental hardware: schematic



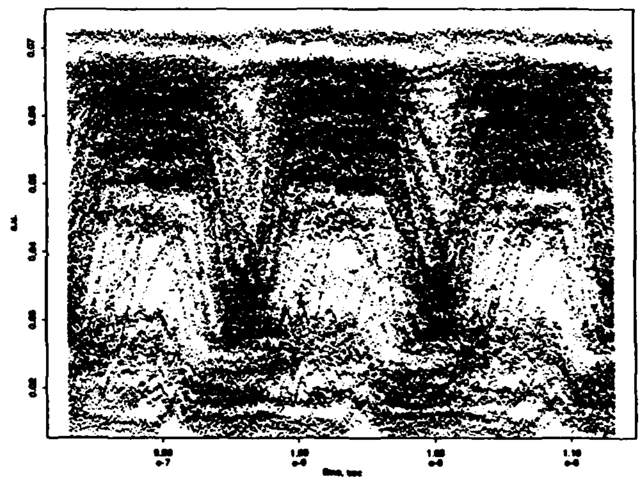
3(a). 32 $POET_1$ outputs at 50 MHz.



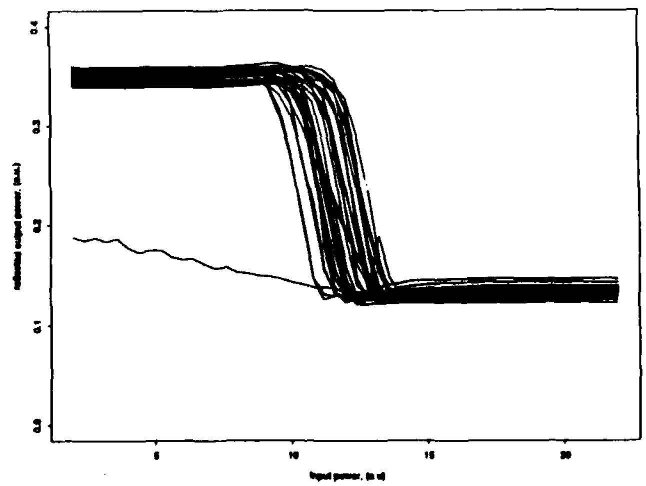
3(b). 128 $POET_1$ outputs at 50 MHz.



4(a). 32 $POET_2$ outputs at 14 MHz.



4(b). 128 $POET_2$ outputs at 14 MHz.



5. Near-DC response of 32 sequentially tested POETs

Design and Construction of Looped Parallel Processors

S. Wakelin and F.A.P. Tooley

Department of Physics, Heriot-Watt University,
Edinburgh. EH14 4AS, U.K.
Fax: 31 451 3136
Tel: 31 451 3027

Introduction.

Recent advances in device technology have indicated that the role of optics in information processing systems will be in providing non-local interconnections between devices, with logic performed by the electronics. Experiments to investigate the viability of different approaches to implementation must identify methods that are extensible for use with more sophisticated devices as they become available

This paper describes some of the design issues arising from the construction of demonstrator systems based on the Cellular Logic Image Processor (CLIP) architecture[1]. This is an application of Single Instruction Multiple Datastream (SIMD) processing. This type of architecture is suitable for implementing optically due to the parallelism of optics and the single instruction mode of the device arrays currently available.

Investigation of optical interconnection implementation must be used to highlight the important issues. The operation of an interconnection module needs to be considered in conjunction with the other parts of the system. The construction of demonstrator systems that perform some function or task, is a valuable way of testing different methods of implementation as well as identifying the critical issues arising from bringing together the various subsystems.

Architecture

Typically an SIMD machine is a computer system consisting of a control unit, N processors, N memory modules and an interconnection network. All the processors are controlled by one unit, and the same instruction is executed by all active processors at the same time. A single instruction stream controls a multiple data stream. The interconnection network provides a means of communicating data between processing elements. This mode of parallel processing is suitable for operations such as image processing and matrix calculations.

The CLIP architecture can be used with non-local interconnections between processing elements in addition to a local interconnection. The general form of the CLIP is shown in figure 1.

Implementation

The implementations described in this paper utilised optical logic devices based on the Self Electrooptic Effect. Symmetric SEEDs (S-SEEDs) [2] operate on the ratio of incident powers, the signal is represented as a differential pair of beams. These devices have been shown to operate well in system demonstrations [3], providing latch or logic functions.

The first SEED implementation of the CLIP architecture used two arrays of S-SEEDs, optically connected in a loop [4]. The signal input to the loop was provided by a nematic liquid crystal spatial light modulator. A local, nearest neighbour interconnection was performed using a binary phase grating. This 1-D fan-out provided two signal inputs from nearest neighbours at the next stage. By performing logic and repeated iterations, it was possible to implement simple image processing algorithms. A simplified schematic of the loop circuit is shown in figure 2.

Further development of the architecture to provide non-local interconnections as well as greater functionality will utilise a module optically implementing a perfect shuffle, and devices that perform logic electronically. This will be a first step towards using smart pixels, that will provide all the necessary gain and logic by the use of integrated transistor technology.

Many of the design issues arising from the implementation of the first system are directly applicable to further developments of the particular architecture chosen, and other types of systems.

Design Issues

Optical design

The widest field point for the 8×16 (300μm×300μm) S-SEED device array used in the CLIP implementation was 1.1° (half field). The array size used was limited by the focussing objective field acceptance, as well as the field curvature contributed by the relay lenses. Diffraction-limited performance for wide field angles is not easily obtained using 'off the shelf' optical components. Custom-designed optical components that can accommodate larger field angles are required for use with large device arrays.

In order to accurately image 1 to 1, attention must be paid to the precision of the effective focal length of the relay lenses. Component fabrication costs increase as the tolerance requirements on the effective focal length increase. A relay lens design that allows small adjustments in focal length when used in a system, ensures that 1:1 imaging is possible. Progress in the development of a suitable relay lens for imaging arrays of beams of greater than 64 × 64 will be reported, in addition to development of a wide, flat-field, f/1 objective. Development of the potential of bulk imaging optics is one area that is essential to the progress of free-space, digital optical information processing systems.

Given that a good optical design is achievable and practical, other optical components need development to perform adequately at such wide fields. In particular, standard beamsplitter cubes that are commercially available only work well at small field angles. Commercially available custom designed coatings for polarising beamsplitter cubes have been used, but the cost makes it viable and attractive to attempt in-house design and coating that could perform better.

Optomechanical development

Progress in the optomechanical design of the S-SEED CLIP and other similarly mounted systems [5,6], has shown that custom-designed mounting techniques are essential to provide the stability and cost-effectiveness required. Clearly, larger array sizes will increase the demands on positional accuracy, for example, some limitations have been noted in rotational alignment of large array BPGs. The fabrication and implementation of the optomechanical components has been constructive in terms of gaining information about the validity of different approaches to performing particular tasks. These problems are often not apparent until the component is actually used in a system. Continual progression of the methods used in implementation is required, as well as new ideas.

Interconnection implementation

The S-SEED CLIP implemented a nearest neighbour interconnection. Basic research into the various methods of performing a Fourier plane fan-out has emphasised which methods of implementation are most practical. It has also drawn attention to some of the disadvantages inherent in the methods used. Further work ideally emphasises the power of optics in providing more global interconnections, thus utilising the main advantages. There has been considerable progress made in reconfigurable devices for interconnection, for example, nematic liquid crystal devices. Dynamic, programmable components could provide practical and flexible methods for interconnection implementation.

Designs of interconnection modules must clearly be implemented as part of a whole system. Independent modules for demonstrating interconnection implementation are valuable as part of early system development, though demonstration as part of a cascadable system is essential.

Conclusions

Progress in free-space digital optical interconnection implementation requires research into all the system components to be carried out in conjunction with each other. Quantitative analysis of the performance of systems and subsystems is essential in order to utilise the advantages that optics could provide as part of a parallel information processing system that combines optics with electronics.

References

[1] M.J.B. Duff and T.J. Fountain, "Cellular Logic Image Processing", Academic Press (1986).

[2] A.L. Lentine, H.S. Hinton, D.A.B. Miller, J.E. Henry, J.E. Cunningham and L.M.F. Chirovsky, "Symmetric SEED, optical set-reset latch, differential modulator/detector", IEEE J. Quant. Electron., **45**, 8 (1989).

[3] F.B. McCormick, F.A.P. Tooley, T.J. Cloonan, J.L. Brubaker, A.L. Lentine, R.L. Morrison, S.J. Hinterlong, M.J. Herron, S.L. Walker and J.M. Sasian, "Experimental investigation of a free-space optical switching network using S-SEEDs" Appl. Opt., **31** (26) p5432 (1992).

[4] F.A.P. Tooley and S. Wakelin, "Design of a S-SEED Cellular Logic Image Processor" Accepted for publication in Appl. Opt. (1992).

[5] F.B. McCormick, F.A.P. Tooley, J.L. Brubaker, J.M. Sasian, T.J. Cloonan, A.L. Lentine, R.L. Morrisson, R.J. Crisci, S.L. Walker, S.J. Hinterlong and M.J. Herron, "Optomechanics of a free-space photonic switch: the system" SPIE Proceedings, **1533** (1991).

[6] J.L.Brubaker, F.B.McCormick, F.A.P.Tooley, J.M.Sasian, T.J.Cloonan, A.L.Lentine, S.J.Hinterlong and M.J.Herron, "Optomechanics of a free-space photonic switch: the components" SPIE Proceedings, **1533** (1991).

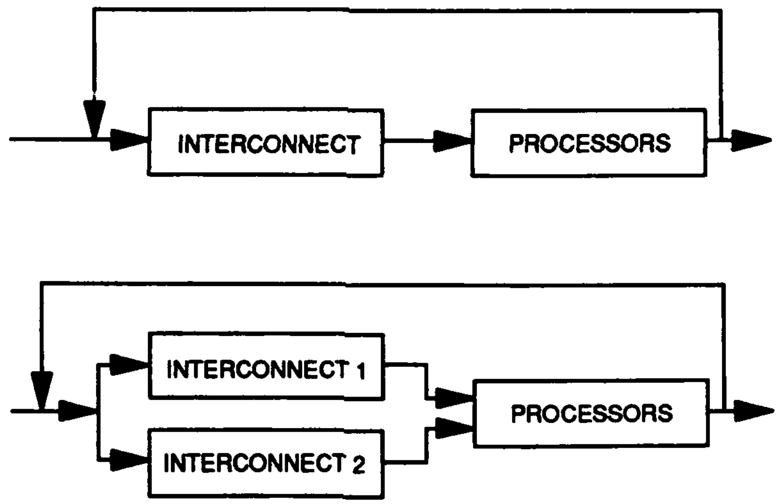


Figure 1. General form of the CLIP architecture

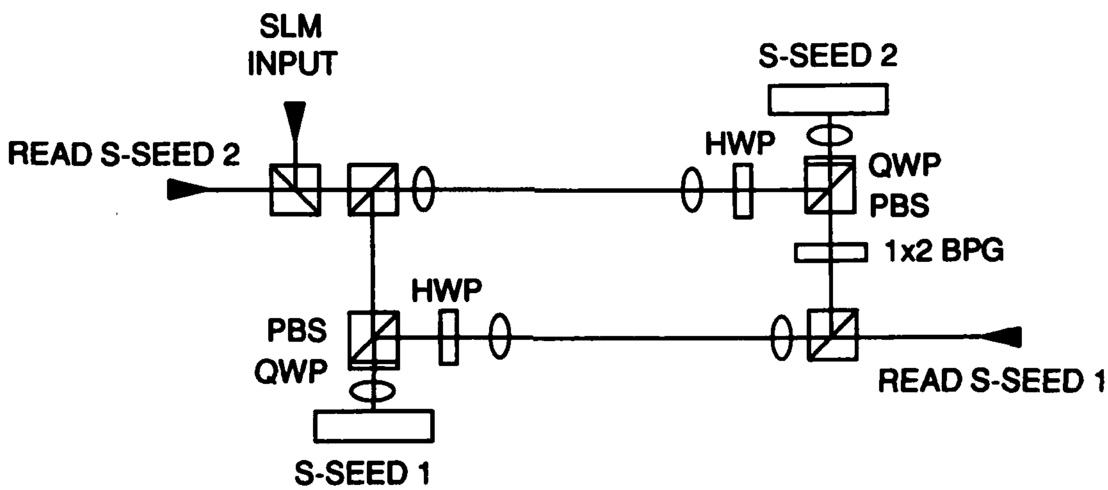


Figure 2. Simplified schematic of S–SEED loop circuit

Smart Pixels: 2

OThB 10:30am-12:00pm
Grand Ballroom West

Joseph W. Goodman, *Presider*
Stanford University

Applications of Smart Pixels

Thomas J. Cloonan
AT&T Bell Laboratories
200 Park Plaza
Naperville, IL 60566

Summary not available at time of printing.

The Application of Self-Similar Patterns to Opto-electronic Shuffle/Exchange Network Design

Michael W. Haney

BDM International, Inc., 7915 Jones Branch Drive
McLean, VA 22102-3396, (703) 848-5854

Introduction/Background

Multistage Interconnection Networks (MINs), based on free-space optical interconnects, were proposed to overcome the communications bottlenecks, skew, and crosstalk associated with long electronic interconnects.[1] Versions based on 2-D arrays of processing elements (PEs) were suggested that more fully utilize the third dimension for higher density and efficient use of optical space bandwidth product (SBWP).[2,3,4] Several of these concepts are based on off-axis imaging techniques which perform shuffle[5] based permutations by optically interleaving equal sized sectors (such as quadrants) of the source array and overlaying the result on an identical array of detectors. The detected signals are then subjected to local exchange/bypass switching elements which operate on small groups of the array and route the signals to the next stage's source array. The various schemes offer trade-offs between the number of stages necessary for an arbitrary permutation and the complexity of the local switching elements. Examples of proposed optical networks include the 2-D separable perfect shuffle,[2] the folded perfect shuffle,[3] and higher order k-shuffles.[4] Figure 1 is an example of a 1-D perfect shuffle, based on off-axis imaging and interleaving of two halves of the array. The figure is also a side view of the 2-D folded perfect shuffle or 2-D separable shuffle, based on the off-axis imaging and interleaving of the four quadrants of the array. Figure 1 depicts the input and output arrays to be identical, as required for cascadability. The optical efficiency for those emitters located at the outer regions of a quadrant can be increased by adding a lens over each quadrant, as shown in the figure, to capture the cone of light from each pixel's emitter and center it on the associated shuffling lens.

Regular Grid Packaging Limits

The density of Optoelectronic Input/Output (OE I/O) circuitry on an OE Integrated Circuit (OEIC) chip is constrained by the heat dissipation requirements of active sources, if used, and chip real estate requirements for the support and logic circuitry at each pixel. Furthermore there may be other considerations, such as the use of multiple OE I/O at each pixel to permit pipelining.[6] The off-axis imaging shuffle interconnects, however, dictate that each pixel must be positioned at a node of a square grid array. Ideally, the circuit designer would be able to position the circuitry across the entire input plane, with the OE I/O on a square grid, with a density limited only by thermal and other logic design considerations. In this case the maximum pixel density would be $1/p^2_{min}$, where p_{min} is the minimum allowed spacing between pixels, as determined by circuit constraints.

For very large arrays (i.e., 32x32 or greater), we must assume that the pixel array will be distributed across many OEICs that will be carefully aligned in the PE plane for optical interconnection. Even the most dense Multi-chip Module (MCM) chip packaging technology will require some separation between chips to allow for electrical connections and heat dissipation. Other practical considerations which limit the density of chips stem from cost issues driven by the need for ease of design, manufacture, test, and repair. For an average high performance application, current MCMs achieve about 40% area efficiency, defined as the ratio of active chip area to package area. Furthermore, this area efficiency is projected not to increase very much in the foreseeable future due to the above practical considerations.[7] With large areas of the plane in which no OE I/O can reside, the result is a lower limit on the density of PEs supported by the package, determined by the minimum spacing between chips on the MCM. Figure 2 illustrates this limit for a regular 5×5 OEIC chip array with minimum chip spacing d. The square grid dot array in Figure 2 represents the closest packing allowed for the source or detector arrays. The highest density PE array is upper bounded by $1/d^2$. For example, if the chip spacing is limited to d=.5 cm, and p_{min} is .1 cm, then the overall packing efficiency is limited to $(.1/.5)^2$ or only .04 of the desired capability. To achieve higher densities from the shuffle techniques described in the references will require tighter packaging between chips (smaller d) or Wafer Scale Integration for the OEIC. The former approach will run into technical and practical packaging limits; the latter will run into reliability and yield constraints.

Self-Similar Grid Approach

The proposed solution is based on the notion of self-similarity of certain fractal sets.[8] A prescription for the application of fractal sets to smart pixel array layout is as follows. The 1-D N pixel perfect shuffle is considered first, then extended to 2-D N^2 arrays, and finally generalized to higher order shuffles. Beginning with a line segment of length 1, remove a central portion of percentage p to leave two equal line segments of length η. From each of these two segments remove a percentage $p=1-2\eta$, and continue this process, successively removing the central $1-2\eta$ portion from each remaining segment. If the process is continued *ad infinitum*, the resulting set of points is a fractal set. These sets are self-similar, that is they are invariant to scale changes of $1/\eta$. To use these sets to form pixel grid patterns for shuffle exchange networks, the formation of the set is terminated at a point where the number of line segments is equal to the number of pixels to be arrayed in one dimension. A pixel location is assigned to the middle of each line sub-segment. Such an assignment is illustrated in Figure 3. Magnification of the array by $1/\eta$ and ignoring one side of it yields an identical array with every other element missing. If another array is also magnified and properly positioned over the first, then an identical array to the original is obtained, just as in the regularly spaced array of Figure 1. This overlaying process is precisely what is required in the shuffle networks. The placement of each pixel of an N element array, with self-similarity parameter η, is given by:

$$x_{pix} = D((1-\eta)/2)(\pm 1 \pm \eta \pm \eta^2 \pm \ldots \pm \eta^{n-1}), \tag{1}$$

where D is the width of the square PE array plane, and $n=\log_2 N$. Note that with $\eta=.5$ the pattern degenerates to a regular, equally spaced array, with spacing D/N. Figure 4 illustrates the side view lens placement necessary to achieve the self-similar array interleave pattern. This technique will have a higher light efficiency, owing to the bunching of OE I/O, and may avoid the need for the auxiliary lenses at the input plane shown in Figure 1. The pair of off-axis imaging lenses that perform the interleaving is placed to provide a magnification of $1/\eta$. From geometrical considerations, their centers are located at positions off of the central axis given by:

$$x_{len} = \pm D((1-\eta)(1-\eta^n)/2(1+\eta)). \tag{2}$$

Note that this scheme naturally places paired pixels in close proximity to each other. This may be advantageous because each pair of signals will be input to the same local exchange/bypass switch. To obtain a 2-D array of N^2 smart pixels with this approach, we simply do the same thing in the orthogonal direction. Figure 5 shows the chip array of Figure 2 with a self-similar grid pattern, generated with the technique described above with $\eta=1/3$, overlaid. In this case the pixel patterns are identical at each chip site except for the 9 chips along the axes, which could be used for all-electronic ICs. The details of the design will be determined by the overall circuit's performance requirements, as well as the need for compatibility with MCM design techniques.

To estimate the improvement in pixel density provided by this approach, we first make the following assumptions. For the perfect shuffle of N^2 PEs (where N is assumed to be a power of 2) we assume there are M^2 identical OEIC chips and P^2 PEs on each chip (M and P are also assumed to be a power of 2). If we assume that η and the square OEIC chip dimensions are calculated such that the closest pixels on an OEIC chip of the array are separated by p_{min}, and the closest OEIC chips are separated by d (typically with $d \gg p_{min}$), then an estimate of the area packing efficiency is derived to be:

$$A = (N(\eta^{n-1})/2p_{min})^2. \tag{3}$$

As a numerical example, consider an N^2=1024 pixel array, distributed across a 4x4 OEIC chip array (M=4), with 8x8 pixels on each chip (P=8). For d=.5 cm and p_{min}=.1cm, the self similar parameter and chip size are derived to be approximately η=.413 and d_{chip}=1.17cm, respectively. The area efficiency, from Equation (3), is calculated to be A=21.6 pixels/cm^2. When compared with the pixel density achievable with a regular square array, $1/d^2$=4 pixels/cm^2, we see an improvement of more than a factor of 5 in OEIC plane area to achieve the same array size. For a given speed of imaging optics, the optical volume required for this example is more than a factor of 10 smaller than the equivalent square grid array—a sizable savings.

The generalization of this approach to higher order shuffles is straight forward. In 1-D, whereas the perfect shuffle follows from dividing and interleaving 2 sectors of the array, a k-shuffle follows from interleaving k equal sized sectors of the array. The 2-D optical implementation of a 2-D separable k-shuffle therefore requires kxk appropriately positioned off-axis imaging lenses. Following the above procedure for the perfect shuffle (or 2-shuffle), we first consider a 1-D array. Figure 6 depicts the array layout procedure for a 4-shuffle-based MIN. The unit line segment is first divided into k equally sized and equally spaced segments, each of which is further similarly divided, and so on. For a k-shuffle, the self similarity parameter, η, must be less than or equal to 1/k. As before we extend the approach to 2-D arrays by using the same pixel placement scheme for the orthogonal dimension. The bunching of OE I/O in the k-shuffle provides improved optical efficiency in the same manner as in the perfect shuffle described previously.

Conclusions

With the rapid advances being made in OE technology, there is increased attention to the issues of packaging, producibility, and design standards. It has been generally assumed that smart pixel arrays should be standardized by distributing the pixels on regular square grid arrays. This paper refutes that notion for the case of optical MIN arrays. The self-similar grid methodology for optoelectronic freespace MIN design overcomes the geometrical limitations to pixel density imposed by high performance MCM IC packaging on square grid arrays. This is achieved by using self-similar grids for the OEIC and pixel patterns that are better matched to the physical and practical constraints stemming from OEIC yield and MCM packaging considerations. Furthermore, the approach offers a better match to the required off-axis imaging optics, thus providing higher optical efficiency across the pixel array. The resulting design should be significantly lower in volume and therefore cost.

References

1. A. Lohmann, W. Stork, G. Stucke, 1985 Opt. Comp. Mtg., WA3.
2. S. H. Lin, T. F. Krile, and J. F. Walkup, *Proc. SPIE*, Vol. 752 (1987).
3. C. Stirk, R. Athale, and M. Haney, *Applied Optics* 27, 202 (1988).
4. A. Sawchuk and I. Glaser, *Proc. SPIE*, Vol. 963 (1988).
5. H. S. Stone, *IEEE Trans. Computers*, C-20, 153-161 (1971).
6. M. W. Haney, *Optics Letters*, Vol. 17, No. 4, February 15, 1992.
7. C. E. Bauer, *Proc. of International Sym. on Microelectronics*, October, 1992.
8. B. Mandelbrot, *The Fractal Geometry of Nature*, Freeman, San Francisco, 1983.

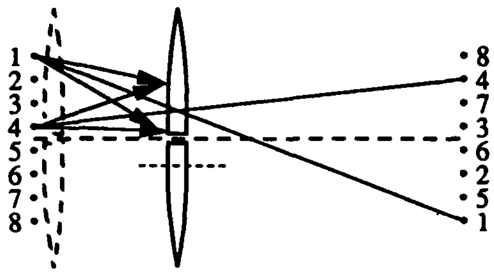


Figure 1. Side view of 2-D separable shuffle or folded perfect shuffle. The auxiliary lenses (dashed) improve optical efficiency by centering each pixel's beam on its shuffling lens.

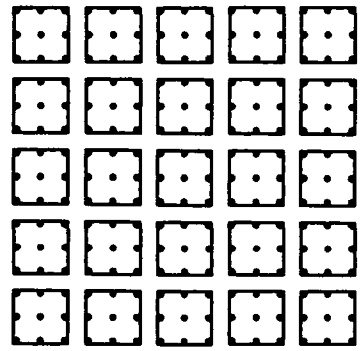


Figure 2. Smart pixel density limitation for MIN arrays constrained to square grid patterns.

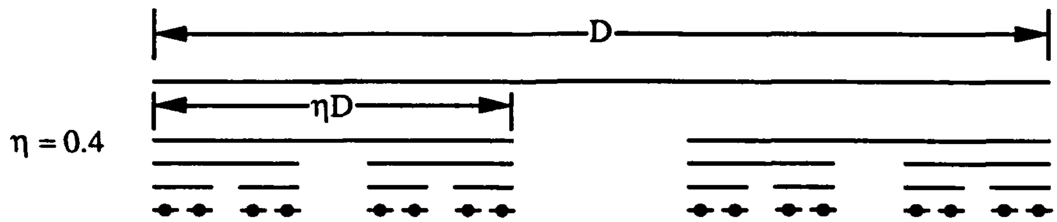


Figure 3. Illustration of pixel OE I/O placement based on a self-similar pattern.

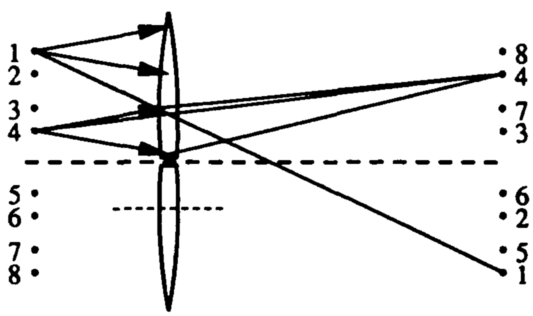


Figure 4. Side view of improved shuffle based on self-similar grid pattern.

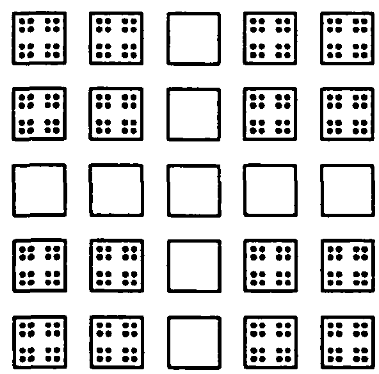


Figure 5. Example of possible 16 x 16 array OEIC configuration with self-similarity parameter $\eta = 1/3$. In this case, the ICs along the axes contain no OE I/O.

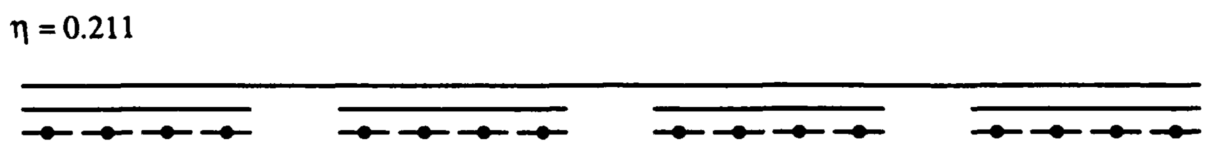


Figure 6. Generalization to 4-shuffle based MIN of self-similar grid approach, showing OE I/O placement for N = 16.

Non-Blocking Optical Interconnection Networks using Electro-Optical Matrix Switches based on VSTEPs

S. Araki, S. Kawai, H. Kurita, T. Numai*, K. Kasahara* and K. Kubota

Opto-Electronics Research Laboratories, NEC Corporation
1-1 Miyazaki 4-chome, Miyamae-ku, Kawasaki 216, Japan
* 34, Miyukigaoka, Tsukuba 305, Japan
Phone: +81-44-856-2107 FAX: +81-44-856-2224

I. Introduction

Optical interconnections have promising features for interconnection networks in massively parallel computers. One of the reasons is that light beams can cross each other with no mutual interference. Particularly promising are free-space optical interconnections, which have no physical pathways and which utilize 3-D connections among processors and switching nodes. We have previously proposed vertical-to-surface transmission electro-photonic devices (VSTEPs) whose functions include light-emission, photo-detection, threshold-switching, and memory[1][2]. Using them, we were successful in producing switching nodes[3] with which it should be possible to achieve not only rigid networks but also reconfigurable interconnection networks in massively parallel computers. Such interconnections require the integration of active devices, such as VSTEPs and vertical cavity surface-emitting laser-diodes, etc., and passive devices, such as microlens arrays. We have proposed various free-space optical interconnection modules which involve VSTEP and microlens arrays[4]. We have used two kinds of microlenses in the modules. One is a planar microlens (PML)[5], and the other is a diffractive optical element fabricated on a VSTEP substrate surface (D-VSTEP)[6]. The modules are expected to provide small-size, high-performance switching nodes and permutation interconnections[7].

This paper describes novel electro-optical (E/O) matrix switches, based on VSTEPs, for use in non-blocking interconnection networks. We first designed full E/O switches, determining on the basis of their optics the maximum number of input channels that we could employ in them, and then, for experimental purposes, designed partial versions of those switches, employing individually addressable VSTEPs with PML arrays. In our experiments, we were able to confirm that the partial versions worked successfully in their elemental functions of addressing and data transmission.

II. Interconnection structures

Figure 1 shows a proposed optical interconnection network scheme, which includes matrix switches and permutation interconnections, and which is based on a 3-stage Clos interconnection network that provides the potential for non-blocking[8].

The matrix switches are of three types: E/O, opto-optical (O/O), opto-electrical (O/E). E/O switches with n-input and m-output (n x m) are used in the first stage. There, electrical data signals received from processors are switched and converted into optical signals, which are then sent through designated permutation interconnections and which are incident to r x r O/O switches in middle stages. m x n O/E switches are used in the last stage to switch and convert received optical signals back into electrical signals and send them out to processors. Each E/O switch is mounted on a board with n processors. All permutation interconnections and O/O switches are integrated into a single module. The E/O switches both switch electrical signals and then convert them into optical signals. They have advantages in switching speed and power consumptions, over a set of a switching device and a converting device.

Total channel number N of this proposed network is defined by the product of the channel number n and r in E/O and O/O switches, as illustrated in Fig.2. 1000 channels networks can be achieved by 32 x m E/O and 32 x 32 O/O switches or 64 x m E/O and 16 x 16 O/O switches. O/O switches are now being developed[4]. In O/O switches input light beams are divided into number of their input-channel. For 32 x 32 switch, the least light power incident to each switching element is -30 dB of the input beam. Then, it is difficult to achieve large channel number switch. We should use E/O switch for enlargement of network channel-number.

III. Optical module design for E/O matrix switch

Figure 3 shows a schematic diagram of the E/O matrix switching system. A node processor sends a data signal to one line devices in a VSTEP array. A controller sends a bit-parallel address signal to the array and that signal selectively turn on the VSTEP devices. Light beams outgoing from the turned on VSTEP devices are propagated and incident to a fiber array by using PMLs and two cylindrical PMLs (C-PMLs). Each fiber receives an optical signal which is sent by each device in one column of a VSTEP array. Therefore, this switch has an electrical-to-optical crossbar connect functions. We may achieve reconfigurable interconnections between node processors and output channels by changing the address signal applied to the E/O switch.

Figure 4 shows an optical module integrated with a VSTEP, a PML and C-PML arrays. In top view, light beam are propagated by the PML and C-PML2. In side view, output light beams from one column devices in the VSTEP array are propagated by the PML array, combined and focused into one of output fibers by two C-PMLs. We evaluated input channel-number n of E/O switches on the basis of their optics. For the purpose, we calculated distance S between PML and C-PML2 by analyzing gaussian beam propagation with ray transfer matrices[4]. The channel-number n was decided by

$$n \leq \frac{D}{p} = \frac{2NA(S_{(p)}+f1)}{p}$$

(p: device pitch, D: C-PML1 width, NA: C-PML1 lens NA, f1: C-PML2 focal length). Figure 5 shows a dependence of their channel-number n on device pitches. The channel number n increases as the device pitch p increase. When we assume that lens NA is equivalent to 0.2, we can decide the device pitch 125 μm and 250 μm for 16 x m and 32 x m E/O switches, respectively. When we use the same size of O/O switches, we may achieve 256 or 1024 channel Clos network. Larger lens NA C-PML1 make it possible to achieve larger channel E/O switches and networks.

IV. Experiments

Figure 6 shows a 1x2 E/O switching experimental setup. For the first step of evaluating the proposed E/O switches, we examined an individually addressable VSTEP array on which a PML array was mounted for examination of free-space beam propagation length and optical data transmission with parallel processors. This structure has a scalability so that it may be easily extended to the structure illustrated in Fig.3. Two VSTEPs in a 4x8 array (Wave length: 950nm, Device pitch: 250 μm, Device size: 10 μm) were used as the optical sources and switching devices. A PML array (NA: 0.22, Focal length: 560μm, Device pitch: 250 μm, Substrate thickness: 1mm) is mounted on the VSTEP array. Two PIN photo-diodes were used for receiving light beams from the VSTEPs. Data signals were amplified up to TTL level by operational amplifiers. The spacing between the PML array and the photo-diodes was 15 mm. The distance is equivalent to 25 input-channels in E/O switch.

Figure 7 shows an applied wave form to the VSTEP and optical data signal which was detected and amplified up to TTL level. In this experiment, switching operation was carried out as follows. Input data signal from node processor and addressing signals from host computer were supplied to VSTEP drivers. Voltage for addressing signal was greater than VSTEP's switching voltage Vs. The drivers added the data and the addressing signal voltages and applied the added voltage to the VSTEP. When the voltage exceeded the VSTEP's threshold voltage which was set with load resistances, the VSTEP emitted optical signals. This data signal was generated by a node processor; Transputers (INMOS T800) whose data transmission speed was 20 Mbps, while VSTEPs and their drivers were operated beyond 100Mbps. Data transmission speed was limited by the node processor. We certified switching operations of individually addressable VSTEPs as fundamental functions of E/O switches.

V. Summary

We proposed a non-blocking optical interconnection which includes E/O, O/O, O/E switches and permutations based on a 3-stage Clos interconnection network. We first designed full E/O switches for use in the network, employing VSTEP, PML and C-PML arrays, determining on the basis of their optics the maximum number of input channels that we could employ in them. For experimental purposes, we designed partial versions of those switches, employing individually addressable VSTEPs with PML arrays. In our experiments, we were able to confirm that the partial versions worked successfully in their elemental functions of addressing and data transmission. We certified the elemental functions of the proposed E/O switches for use in the non-blocking optical interconnection network.

Acknowledgements

The authors gratefully thank M. Oikawa of Nippon Sheet Glass Co., Ltd. for his supplying the planar microlenses used in these experiments. They also thank T. Nagahori for his valuable comments, and are grateful to K. Kobayashi, K. Hotta and K. Asakawa for their suggestions and encouragement.

References

[1]K. Kasahara, *et al.*, Appl. Phys. Lett. **52**, 679-681(1988).
[2]T. Numai, *et al.*, Appl. Phys. Lett. **58**, 1250-1252(1991).
[3]I. Ogura, *et al.*, Appl. Phys. Lett. **57**, 540-542(1990).
[4]S. Kawai, *et al.*, to be published in Proc. SPIE.
[5]M. Oikawa, *et al.*, Proc. SPIE **898**, 3-11(1988).
[6]S. Kawai, *et al.*, Topical Meeting on Smart Pixels, Digest, (1992).
[7]S. Kawai, IEEE J. Lightwave Tech. 9, 1774-1779(1991).
[8]C. Clos, Bell Sys. Tech. J. **32**, 406-424(1953).

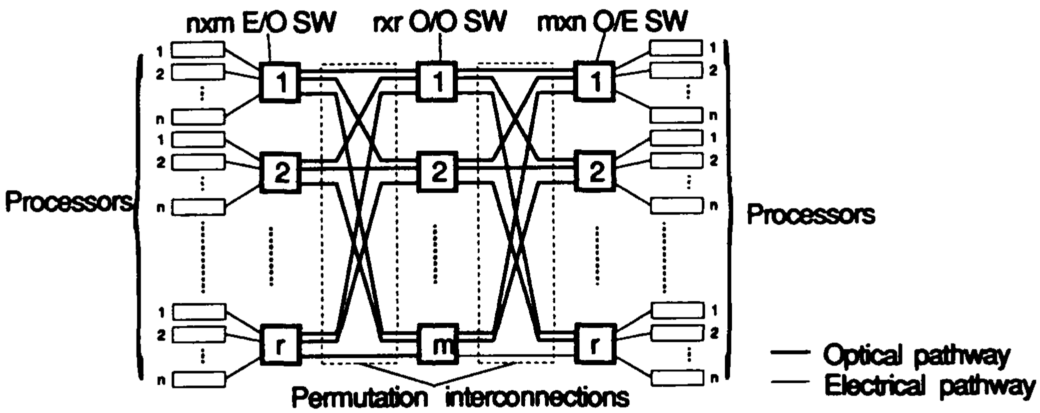


Fig.1 Proposed optical interconnection network.

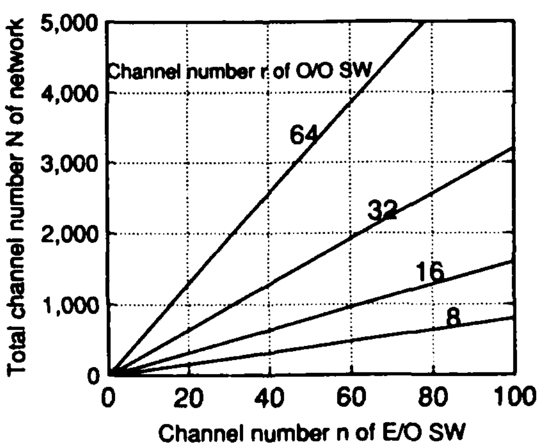


Fig.2 Total channel number of the proposed networks.

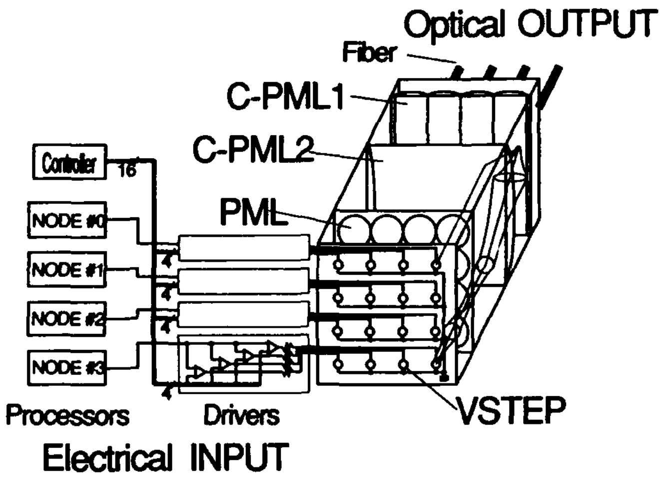


Fig.3 Schematic diagram of E/O matrix switching system.

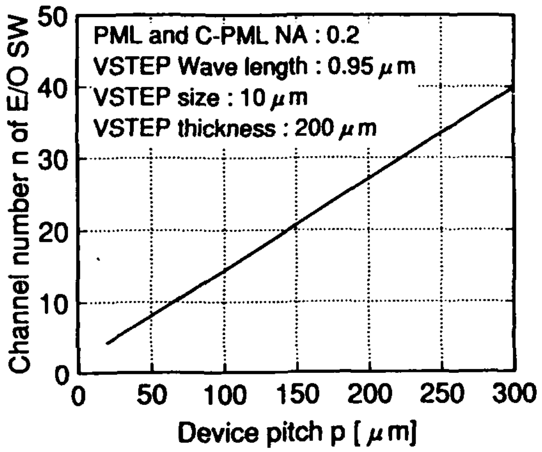


Fig.5 Dependence of E/O switching module channel number on device pitch, when lens NA is equivalent to 0.2.

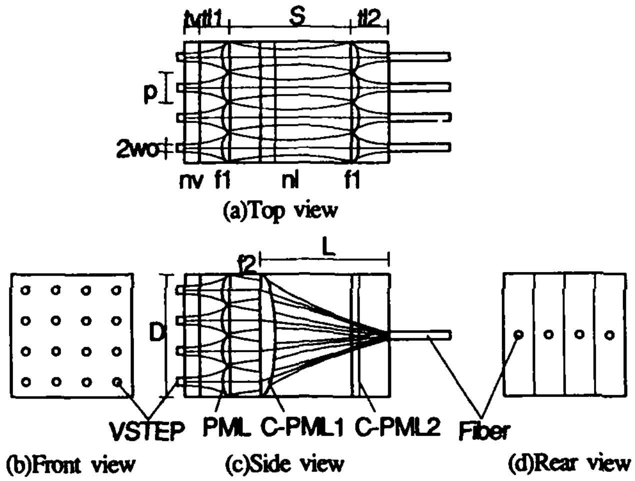


Fig.4 Optical design of E/O switching module.

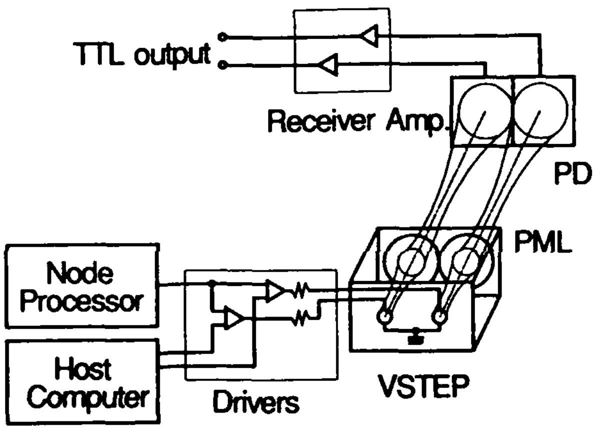


Fig.6 E/O switching experimental set up.

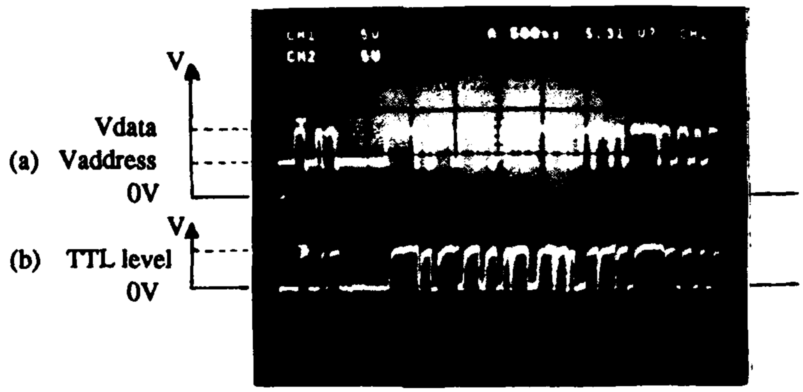


Fig.7 Signal waveforms
(a) Input signal to VSTEP
(b) Output signal from receiver amp.

Packaging considerations for planar optical systems

Bruno Acklin and Jürgen Jahns
AT&T Bell Laboratories
Crawfords Corner Road
Holmdel, NJ 07733

Various integration techniques have been suggested in order to build compact opto-electronic systems [1-3]. For an actual implementation, several topics have to be addressed:

(a) the design of the optical system,
(b) its manufacturability with standard fabrication techniques,
(c) hybrid integration of the passive optics with opto-electronic device arrays,
(d) the thermal management of the system.

To discuss these issues, we consider the model of an optical multi-chip module (MCM), that is based on the planar optics approach [2] (Figure 1).

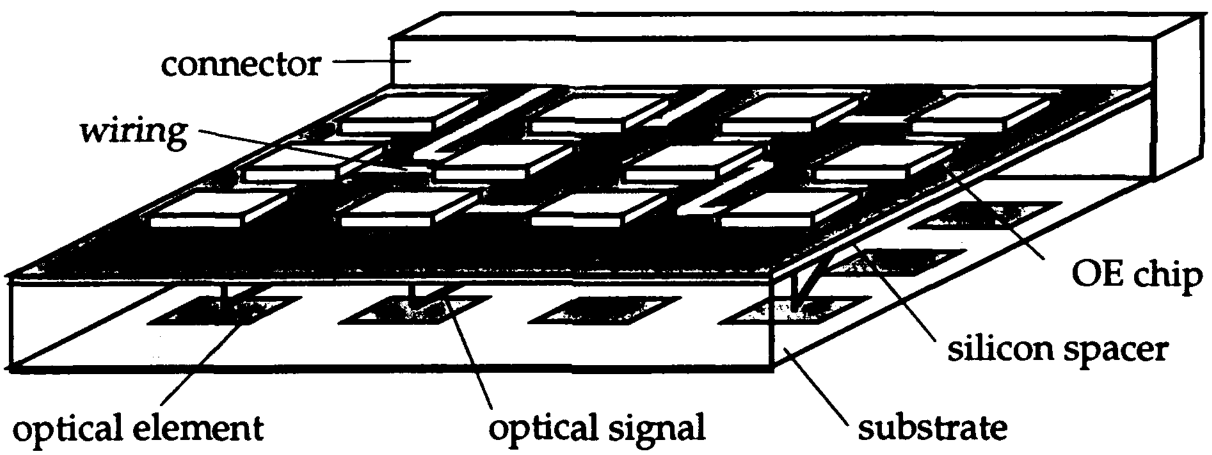


Fig. 1: *Device arrays are placed on an optical substrate in a planar configuration. Light signals travel through the substrate to provide parallel interconnections between the devices.*

Physically, the MCM consists of three functional layers: an opto-electronic layer, a silicon spacer, and the optical backplane (Figure 2). Optical signals are generated and detected by opto-electronic chips. These active chips are mounted on a silicon spacer wafer with windows to allow for light propagation between the active layer and the backplane. The silicon wafer can also help to dissipate heat and may host auxiliary electronics such as drivers. The Si-wafer is in contact with a planar optical substrate, in which the interconnects are implemented using diffractive and refractive microoptical elements, and metallic (or TIR-) mirrors.

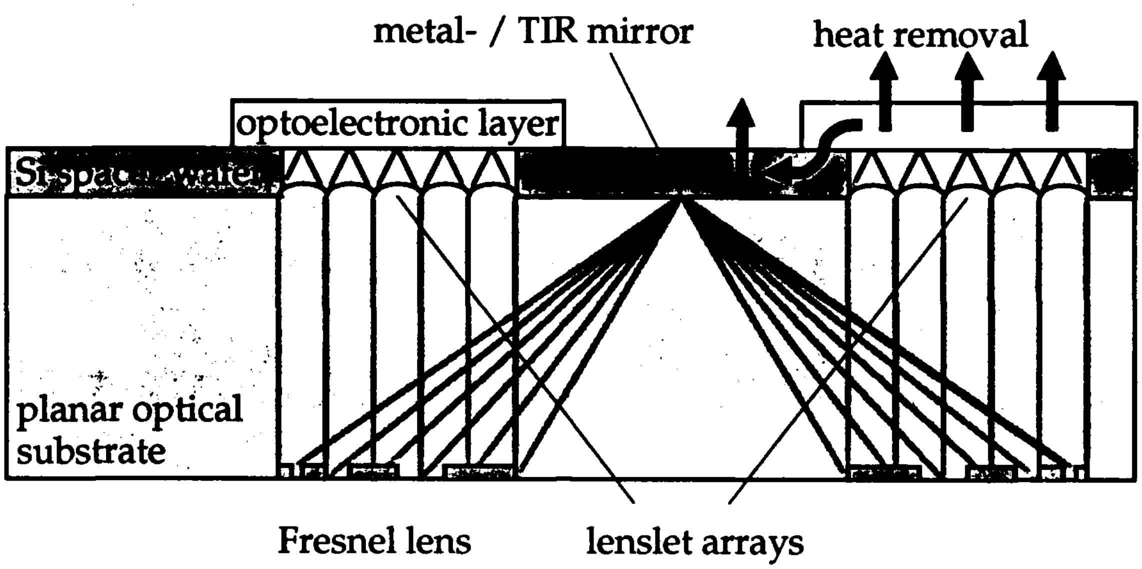


Fig. 2: *Opto-electronic MCM implementation using three functional layers: opto-electronic device layer, Si spacer, and planar optical substrate.*

We are going to discuss the topics mentioned above:

(a) Three generic imaging systems were discussed and compared in a recent article [4]. They are the conventional optical 4f-imaging system, a system based on the use of microlens arrays and a hybrid system that combines the conventional and the microlens approach. The two former systems were already demonstrated in a planar optical implementation [5-6]. Here, we focus on the third approach which is attractive because it allows one to realize a large number of interconnections in a planar optics configuration (cf Figure 2).

(b) The manufacturing of microoptic elements with standard VLSI techniques has been reported by many authors in the past few years. More advanced techniques such as direct electron beam lithography for the fabrication [7] and the sol-gel process for the replication [8] of microoptic elements were also demonstrated successfully.

(c) The hybrid integration of opto-electronic chips onto planar optical substrates has been demonstrated recently [9] using flip-chip solder bump bonding [10]. This technique is well developed and has been investigated for opto-electronic applications earlier [11-12]. Thermal anodic bonding [13] can be used to bond the Si-wafer on the planar glass substrate.

(d) For opto-electronic device arrays that support massive parallel interconnections heat dissipation becomes a major issue. Promising devices are arrays of vertical-cavity surface emitting microlasers [14], and self-electro-optic effect devices (SEED), where arrays with up to 32,000 elements have been demonstrated [15]. Assuming an array of 32 x 32 devices on a mm^2 running at high speed, the dissipated heat will be on the order of kW/cm^2. Conventional cooling techniques such as forced air cooling are able to remove heat amounts of up to 1 W/cm^2 and are therefore insufficient. On the other hand, heat in excess of 1 kW/cm^2 has been removed from 1 cm^2 chips with more sophisticated liquid cooling techniques, based on micromachined cooling channels [16]. We present an estimation of the cooling potential of the Si-wafer, using conduction in the Si itself, or in metal pads, or synthetic diamond layers [17] deposited thereon.

References

[1] Iga, K., Oikawa, M., Misawa, S., Banno, J., and Kokubun, Y., "Stacked planar optics: an application of the planar microlens," *Appl. Opt.* **21** (1982) 3456-3460.

[2] Jahns, J. and Huang, A., "Planar integration of free-space optical components," *Appl. Opt.* **28** (1989) 1602-1605.

[3] Brenner, K.-H., "3D-Integration of digital optical systems," in *Optical Computing*, 1991, Technical Digest Series (Optical Society of America, Washington, DC 1988) **6** (1991) 25-28.

[4] Lohmann, A. W., "Image formation of dilute arrays for optical information processing," *Opt. Comm.* **86** (1991) 365-370.

[5] Jahns, J., "Integrated optical imaging system," *Appl. Opt.* **29** (1990) 1998.

[6] Jahns, J. and Däschner, W., "Integrated free-space optical permutation network," in *Optical Computing*, 1991, Technical Digest Series (Optical Society of America, Washington, DC 1988) **6** (1991) 29-32.

[7] Shiono, T. and Ogawa, H., "Diffraction-limited blazed reflection diffractive microlens for oblique incidence fabricated by electron-beam lithography," *Appl. Opt.* **30** (1991) 3643-3649.

[8] Noguès, J. L., "Pure silica diffraction gratings and microlens arrays made by the sol-gel process," in *Diffractive Optics: Design, Fabrication, and Applications*, (Optical Society of America, Washington, D.C., 1992), **9** , 15-17.

[9] Jahns, J., Morgan, R. A., Nguyen, H. N., Walker, J. A., Walker, S. J., and Wong, Y. M., "Hybrid integration of surface-emitting microlaser chip and planar optics substrate for interconnection applications," *IEEE Phot. Tech. Lett.* **4** Dec. (1992).

[10] Miller, L. F., "Controlled collapse reflow chip joining," *IBM J. Res. and Dev.* **13** (1969) 239-250.

[11] Goodwin, M. J., Moseley, A. J., Kearley, M. Q., Morris, R. C., Kirkby, C. J. G., Thompson, J., Goodfellow, R. C., and Bennion, I., "Optoelectronic components arrays for optical interconnection of circuits and subsystems," *J. Lightwave Technol.* **9** (1991) 1639-1645.

[12] Lehmen, A. v., Banwell, T., Chan, W., Orenstein, M., Wullert, J., Maeda, M., Chang-Hasnain, C., Stoffel, N., Florez, L., and Harbison, J., "Electronically and optically controllable vertical cavity surface emitting microlaser arrays for optical interconnect and signal processing applications," *Proc. SPIE* **1582** (1991) 83-90.

[13] Wallis, G. and Pomerantz, D. I., "Field assisted glass-metal sealing," *J. Appl. Phys.* **40** (1969) 3946-3949.

[14] Jewell, J. L., Lee, Y. H., Scherer, A., McCall, S. L., Olsson, N. A., Harbison, J. P., and Florez, L. T., "Surface-emitting microlasers for photonic switching and interchip connections," *Opt. Eng.* **29** (1990) 210-214.

[15] Chirovsky, L. M. F., Focht, M. W., Freund, J. M., Guth, G. D., Leibenguth, R. E., Przybylek, G. J., Smith, L. E., D'Asaro, L. A., Lentine, A. L., Novotny, R. A., and Buchholz, D. B. , "Large arrays of symmetric self-electro optic effect devices," in *OSA Proc. on Photonic Switching, H. S. Hinton and J. W. Goodman, eds.,* (Optical Society of America, Washington, DC) **8** (1991) 56-59.

[16] Tuckerman, D. B. and Pease, R. F. W., "High-performance heat sinking for VLSI," *IEEE El. Dev. Lett* **EDL-2** (1981) 126-129.

[17] Graebner, J. E., Jin, S., Kammlott, G. W., Herb, J. A., and Gardinier, C. F., "Unusually high thermal conductivity in diamond films," *Appl. Phys. Lett.* **60** (1992) 1576-78.

Optical Interconnections: 1

OThC 2:00pm-3:30pm
Grand Ballroom West

Mark Bendett, *Presider*
IMRA America

Free-space Optical Switching Modules

Masayasu Yamaguchi and Ken-ichi Yukimatsu
NTT Communication Switching Laboratories
3-9-11, Midori-cho, Musashino-shi, Tokyo 180, Japan
Tel: +81-422-59-2306 Fax: +81-422-59-2473

INTRODUCTION

Free-space optical switches are promising candidates for high-density large-scale switches required in optical computing and in telecommunication switching. This is because they use spatial parallelism, an inherent feature of optical technology, which offers high-density interconnections. Various types of free-space optical switches have been studied using liquid-crystal devices, semiconductor devices and so on [1]-[7]. They can be classified as either analog or digital switches. Analog switches maintain the waveform of the signal during switching. Therefore, they are transparent to signal format, bit rate, and modulation type, so they can handle high-speed optical signals over a wide range of wavelengths. Their disadvantage, however, is that they accumulate loss and crosstalk when they are used in a multistage structure. Digital switches, on the other hand, can regenerate the shape of the signal waveform (level and/or timing), so there is no accumulation of loss or crosstalk even in multistage structures. However, the signal speed is limited by the response time of the switching devices. In addition, when the device has multiple-quantum-well structures or cavities, the wavelength range is also limited. Each type of free-space switch therefore has its own range of applications based on these limitations.

Here, we discuss applications of analog and digital free-space optical switches in telecommunication systems and present our recent experimental switches based on liquid-crystal and semiconductor array devices.

APPLICATIONS OF ANALOG SWITCHES

Figure 1 shows two examples of analog-switch applications: a subscriber-line concentrator and an inter-module connector. The concentrator, which connects only live input lines to vacant output lines, is used for economically increasing the number of available input lines in a switching system. The inter-module connector exchanges the optical interconnections between functional modules in the switching system. This exchange of interconnections is necessary for changing the system configuration when one of the modules becomes out of order.

For these applications, analog switches are better than digital switches, because their transparency to signal format, bit rate, modulation, and wavelength makes the switching system more flexible. In addition, these applications do not need high-speed switching because the concentrator changes its connections only at the origination and termination of calls and the inter-module connector changes its connections only when a module is out of order. These applications only require modest switching speeds, such as a few hundred milliseconds, so relatively slow elements like liquid-crystal switches can be used.

APPLICATIONS OF DIGITAL SWITCHES

Figure 2 shows an example of a digital-switch application. Digital switches are suitable for the multistage space-division switches of time-division circuit-switching or packet-switching systems, because they are fast enough for the time-slot or cell switching. We can use GaAs-based array devices as these switches if we convert the wavelength of the optical fiber transmission lines into the wavelength of the switches by using wavelength converters. In addition, if we can make large-scale optical memory arrays, we can also use them for time switches or cell buffers.

EXPERIMENTAL SWITCHES

(1) Call-monitoring concentrator

We have demonstrated an optical concentrator, a multistage analog switch, consisting of liquid-crystal-based beam shifter modules and a light-transmitting two-dimensional photodetector array module for call monitoring (Fig. 3) [6]-[9]. The experimental concentrator had 1024 inputs and 256 outputs, and its dimensions were 15 cm x 19 cm x 27 cm. The total insertion loss of the 8-stage switch part was less than 10 dB.

(2) Holographic switch

A holographic switch is a single-stage analog switch, which can be non-blocking. 1-input-64-output switching and 2-input-32-output switching was experimentally demonstrated using a liquid-crystal display as the hologram recording medium (Fig. 4) [5]. The holographic switch is applicable to non-blocking large-scale switches such as an inter-module connector.

(3) Digital switch

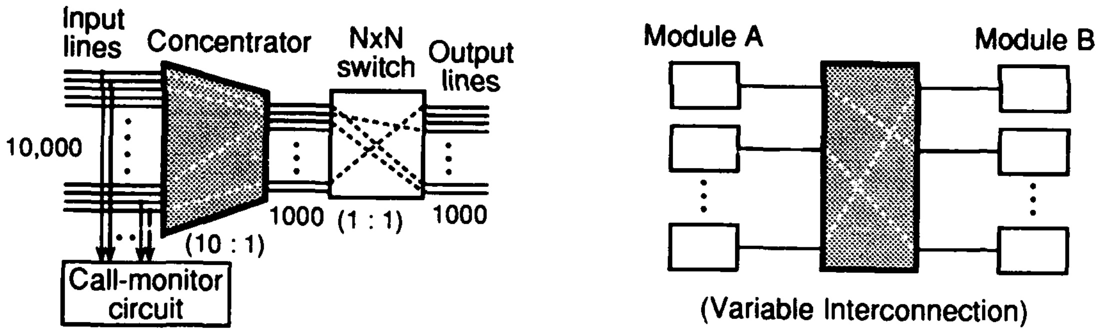
We have fabricated a digital switch (Fig. 5) using a two-dimensional array of exciton absorption reflection switches (EARSs) [10]. The experimental switch demonstrated 2-input-2-output switching using video signals (1.5 Mb/s, CMI code, moving picture). The dimensions of the switch were about 5 mm x 5 mm x 25 mm for a single stage. This type of switch may be applicable to space-division switches in time-division circuit-switching or packet-switching systems.

CONCLUSION

This paper presented our recent analog and digital free-space optical switches and discussed their applications. These types of free-space switches still need further studies before they can be applied to practical systems. For example, the assembly technique must be established as well as the device and system architectures.

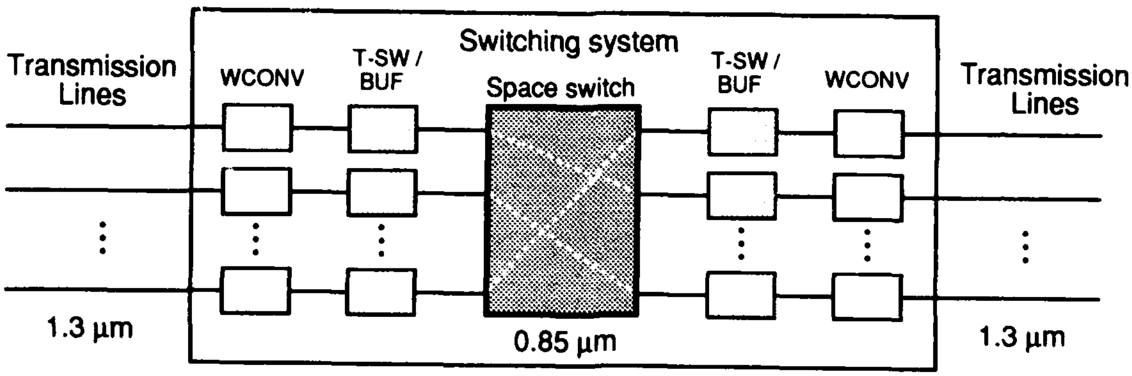
REFERENCES

[1] K. Noguchi et al.: ECOC'91, TuC3-2, Paris, 1991
[2] T. Yamamoto et al.: OEC'92, 17A4-3, Makuhari Messe, 1992
[3] M. Nishio et al.: ISS'92, B10.4, Yokohama, 1992
[4] S. J. Hinterlong: ISS'92, P20, Yokohama, 1992
[5] H. Yamazaki et al.: Opt. Lett., 17, 17, pp. 1228-1230, 1992
[6] M. Yamaguchi et al.: PS'92, 1A5, Minsk, 1992
[7] M. Yamaguchi et al.: OEC'92, PD-9, Makuhari Messe, 1992
[8] S. Shirai et al.: PS'92, 2A6, Minsk, 1992
[9] M. Okamura et al.: OEC'92, 16C4-5, Makuhari Messe, 1992
[10] C. Amano et al.: IEEE Photon. Technol. Lett., 3, 8, pp. 736-738, 1991



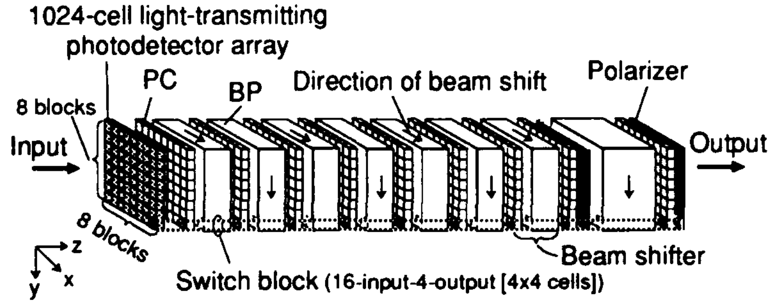
(a) Subscriber-line concentrator (b) Inter-module connector

Fig. 1. Examples of analog-switch applications



WCONV: Wavelength converter T-SW: Time switch BUF: Cell buffer

Fig. 2. Example of a digital-switch application



PC: Polarization controller array BP: Birefringent plate

Input ports	= (8 x 8)blocks x (4 x 4)cells=1024
Output ports	= (8 x 8)blocks x (2 x 2)cells=256

Fig. 3. Structure of the experimental concentrator using cascaded beam shifter modules.

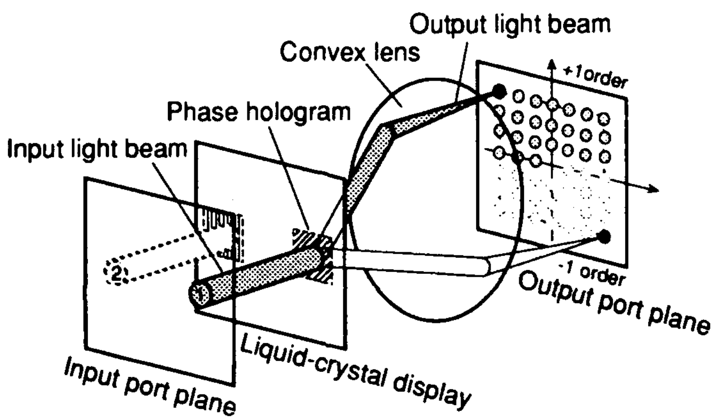


Fig. 4. Structure of the experimental holographic switch using a liquid-crystal display.

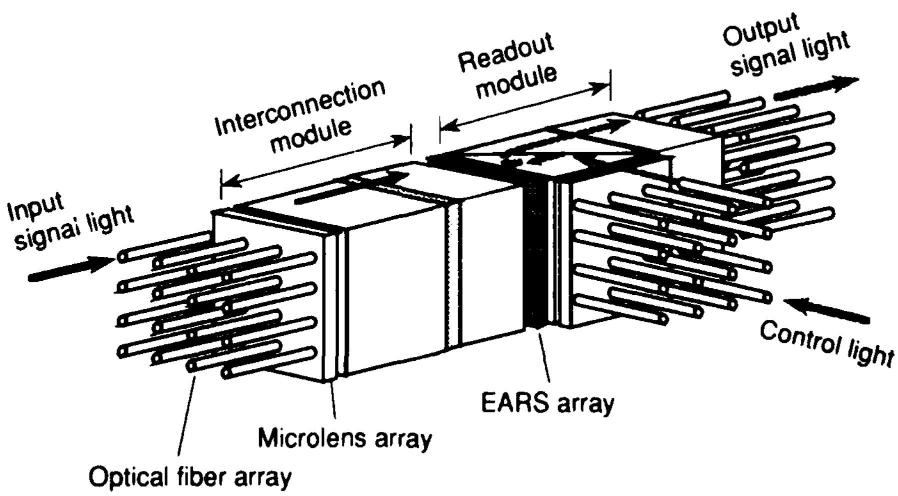


Fig. 5. Structure of the experimental digital switch using an EARS array.

Design issues for beam array generation gratings

Rick L. Morrison and Sonya L. Walker
AT&T Bell Laboratories
263 Shuman Blvd.
Naperville, IL 60566

Introduction

Diffractive optical elements play an important role in digital free-space photonic systems. As phase gratings they provide the functionality of holograms that can not be readily achieved using conventional optics. Two areas that are well served by diffractive elements are the generation of large beam arrays to illuminate an array of opto-electronic devices and beam generation to form the linking stage in the interconnection network[1]. The unique use of the phase grating at the interconnection stage provides an exceptionally compact and straight-forward means of implementing the optical link.

Historically, binary phase grating designs (or Dammann gratings[2]) have progressed from one-dimensional designs generating a small, odd-number of uniform intensity spots to large, two-dimensional even-numbered arrays[3] with arbitrary spot intensities. Three criteria determine the quality of the grating. First, the spot pattern must match the layout of the photonic device array. Second, the relative intensities of the fabricated gratings must meet design tolerances. Finally, light should be efficiently coupled into the designated orders.

The issues influencing diffractive grating designs for free-space photonic systems include:

Design The spot array format should match the photonic device array format. The relative spot intensities must either be uniform or meet other specifications. High efficiency is desirable.

Fabrication Limitations of the fabrication process (microlithography and reactive ion etching) influence actual performance.

Operation Wavelength control and stability of the light source is critical for large array designs to ensure correct registration of the spots with the device array. Characterization systems are necessary to measure how well operation conforms to design expectations.

In the future, some free-space digital optical systems will utilize smart pixel devices[4] that thereby place further demands on the supporting optical infrastructure. These smart pixel system will resemble previous photonic systems[5,6], in that information will be transferred via light beams. However, a much larger fraction of the device substrate will be devoted to electronic processing. The additional functionality leads to larger separations between sets of optical communication ports associated with each electronic cell. A larger spot separation results in a smaller grating period and results in tighter tolerances on the fabrication process.

Design

The success of even-numbered array designs was determined in part by the use of even-numbered device arrays and on the desire to eliminate the large intensity non-uniformity introduced by the central order beam. Gratings have been demonstrated that generate ten of thousands of beams. The design of grating patterns have shifted from the standard rectangular cells of the moderately efficient binary phase grating, to arbitrarily shaped binary phase cells composed of non-separable two-dimensional solutions[7], to multilevel and continuous phase[8] approaches.

The format of the spot array has expanded beyond the early requirements of a simple regular array of uniform intensity spots. For example, when gratings are used as to form the interconnection stage for a Banyan network[1], the suppression of various noise orders outside of the uniform intensity region becomes a critical concern[9]. These extraneous orders must be suppressed since they could fall on alternate receivers and thereby corrupt the data routed to that location. For example, figure 1 shows the region of uniform intensity spots, the region of suppressed orders, and an external, unspecified intensity region for one grating design. The addition of suppressed orders to the design process is accomplished by adding criteria describing their reduced intensity to the cost parameter in the optimization program used to determine the grating pattern.

Typical smart pixel systems will most likely continue to use dual differentially encoded pairs of light beams to transmit information. It will be advantageous for high speed operation to keep the two modulators closely spaced. This configuration leads to a spot array with two characteristic spot separations, the intercellular periodicity of hundred's of microns and the modulator pair separation of tens of microns. An example is shown in figure 2. A solution to this problem is the extension of suppressed order designs to internal orders within the spot array. In the figure, lighter dots represent suppressed intensity orders lying between the usual uniform intensity spots. In this manner, an irregularly spaced set of spots is designed from a set of regularly spaced orders.

Fabrication Issues

During fabrication, the phase design must be accurately replicated onto either a transparent or reflective substrate using microlithographic techniques. This means that the relative location of phase transitions must be accurately transferred and that the precise depth of material be removed during reactive ion etching. The phase transition can typically be held to about a 0.1μm accuracy during fabrication of a binary phase design.

An implication of smart pixel system design is that spot separation is usually about an order of magnitude larger than in previous systems. This larger spacing results in significantly smaller grating periods when comparable optics are employed. For the smart pixel device array cited earlier, the period would be closer to 130μm compared to 1300μm typical of previous demonstration systems. This period shrinkage increases the complexity of faithfully reproducing the design on the substrate via lithography due to the limited resolution available in the mask production step. This quantization of the spatial features and the pattern period leads to intensity variations.

Period quantization problems become evident in the design of the holographic interconnection gratings linking arrays requiring large separations between interconnected devices. Based on an 850nm wavelength illumination, a 15.61mm focal length objectives, and a 480μm order spacing for a 3 spot grating, a period value of 27.64μm is required. Typical quantization of electron beam mask writing is only 0.1μm, however. For this simple design, the quantization of the period is substantially more severe than the quantization of the pattern features. Fortunately, this problem can be addressed and the solution will be described in forthcoming paper.

The need for higher efficiency coupled with the decreased size of the grating period size can lead to poorer performance due to fabrication limitations. Figure 3 shows the dependence of the spot intensity uniformity on alignment accuracy between the two masks required to fabricate a simple 4 level grating that generates a line of four spots. The 120.4μm period design produces a spot separation of 210μm when used with a 15.61mm objective lens. The graph shows that a placement accuracy of 0.25μm still causes a intensity spread of 5%.

Therefore, more complicated multi-level designs will be expected to have significantly worse intensity uniformities compared to comparable binary phase designs. One alternative approach is to use nonseparable binary phase designs based on polygons having efficiencies intermediate between those of the standard rectangular binary phase designs and multilevel designs. Another approach is to eliminate the multi-step lithography process by directly writing a continuous phase profile into photoresist.

Operational Characteristics

Wavelength sensitivity is a critical issue for diffractive systems that operate with semiconductor laser diodes. For example, the illumination of a large 64x32 S-SEED device array or a moderate size array of 4x4 smart pixels each having a spot alignment tolerance of about 1μm requires a wavelength control of about 1-2nm at 850nm. Such an accuracy is difficult to maintain due to modal instabilities of high power laser diodes resulting from reflection from components and modulation. The required accuracy will become tighter as the size of the device array increases. One solution to this problem is to incorporate a feedback mechanism, via a grating or a Fabry-Perot cavity within the laser system to stabilize and control the wavelength. Another possibility is to use a diode pumped solid state laser, although currently no adequate source exists near 850nm.

The characterization of beam array generating gratings is necessary to determine if system tolerances are fulfilled and to aid in improving the fabrication process. We have developed a CCD based system that can measure the operation of large spot arrays with an accuracy of a few percent for the intensity. This system also proves valuable in tracking the accumulation of intensity non-uniformity due to optical aberrations that accumulates as spot arrays are transported through a photonic system. Figure 4 shows the measured intensity distribution of a 64x32 grating.

Summary

We have described several issues concerning spot array generation gratings for current photonic systems and future smart pixel based systems. The application of suppressed orders in the grating design allow configurations that appear irregularly spaced or require suppressed intensities to eliminate causes of signal interference. Unfortunately, the large size of the smart pixel cell leads to small grating periods that results in quantization problems and greater sensitivity to alignment error for multilevel designs.

References

[1] R.L. Morrison, S.L. Walker, T.J. Cloonan, F.A.P. Tooley, F.B. McCormick, J.M. Sasian, A.L. Lentine, J.L. Brubaker, R.J. Crisci, S.J. Hinterlong, and H.S. Hinton, "Diffractive optics in a free-space digital optic system," in Diffractive Optics: Design, Fabrication and Applications Technical Digest, 1992, Vol. 9, pp. 28-30.

[2] H. Dammann and K. Gortler, "High efficiency in-line multiple imaging by means of multiple phase holograms," Opt. Comm. Vol. 3, 312-315 (1971).

[3] R.L. Morrison and S.L. Walker, "Progress in diffractive phase gratings used for spot array generators," in Optic Computing 1991 Technical Digest, Vol. 6, 144-147 (1991).

[4] A.L. Lentine, T.J. Cloonan, H.S. Hinton, L.M.F. Chirovsky, L.A. D'Asaro, E.J. Laskowksi, S.S. Pei, M.W. Focht, J.M. Freund, G.D. Guth, R.E. Leibenguth, L.E. Smith, G.D. Boyd, and T.K. Woodward, "4x4 arrays of FET-SEED embedded control 2x1 switching nodes," in the IEEE Topical Meeting on Smart Pixels, Postdeadline papers, 1992.

[5] F.B. McCormick, F.A.P. Tooley, T.J. Cloonan, J.L. Brubaker, A.L. Lentine, R.L. Morrison, S.J. Hinterlong, M.J. Herron, S.L. Walker, and J.M. Sasian, "Experimental investigation of a free-space optical switching network by using symmetric self-electro-optic devices," Appl. Opt. Vol 31, No. 26, 5431-5446 (1992).

[6] F.B. McCormick, F.A.P. Tooley, J.L. Brubaker, J.M. Sasian, T.J. Cloonan, A.L. Lentine, R.L. Morrison, R.J. Crisci, S.L. Walker, S.J. Hinterlong, M.J. Herron, "Optomechanics of a free-space photonic switching system fabric: the system," Proc. of the SPIE symposium on Optomechanics and Dimensional Stability, vol. 1533, pp. 97-114 (1991).

[7] A. Vasara, M. Taghizadeh, J. Turunen, J. Westerholm, E. Noponen, H. Ichikawa, J.M. Miller, T. Jaakkola, and S. Kuisma, "Binary surface-relief gratings for array illuminators in digital optics," Appl. Opt. vol. 31, no. 17, pp. 3320-3336 (1992).

[8] H.P. Herzig, D. Prongue and R. Dandliker, "Design and fabrication of highly efficient fan-out elements," Japan. Jour. of Appl. Phys. vol. 29, no. 27, pp. 1307-1309 (1990).

[9] H. Heibmeier, U. Krackhardt, and N. Streibl, "A Dammann grating with diffraction order of arbitrary intensity etched into Al_2O_3," Opt. Comm, vol. 76, no. 2, pp. 103-106 (1990).

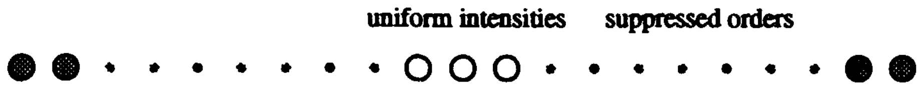


Figure 1: Spot array containing noise orders, suppressed intensity orders and 3 uniform intensity orders.

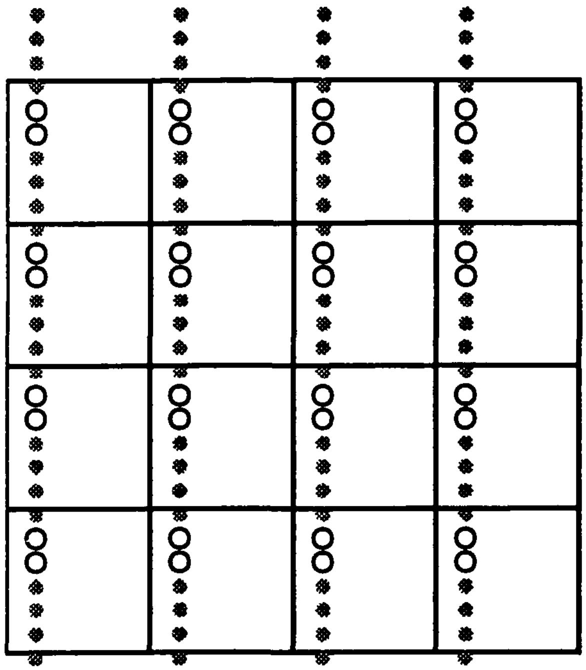


Figure 2: Irregularly spaced spot array with internal suppressed intensity orders.

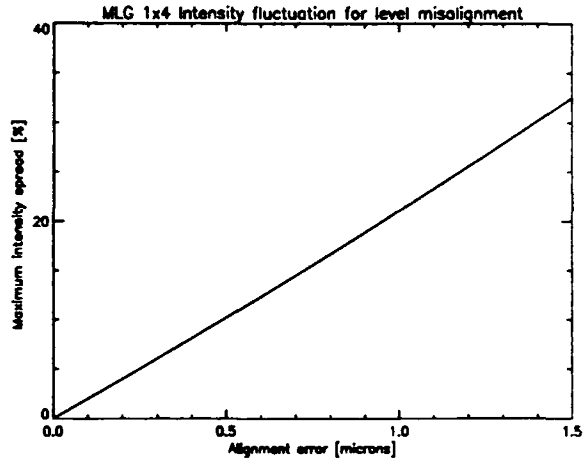


Figure 3: Uniformity vs. alignment for a 4 level 1x4 design with 120 micron period.

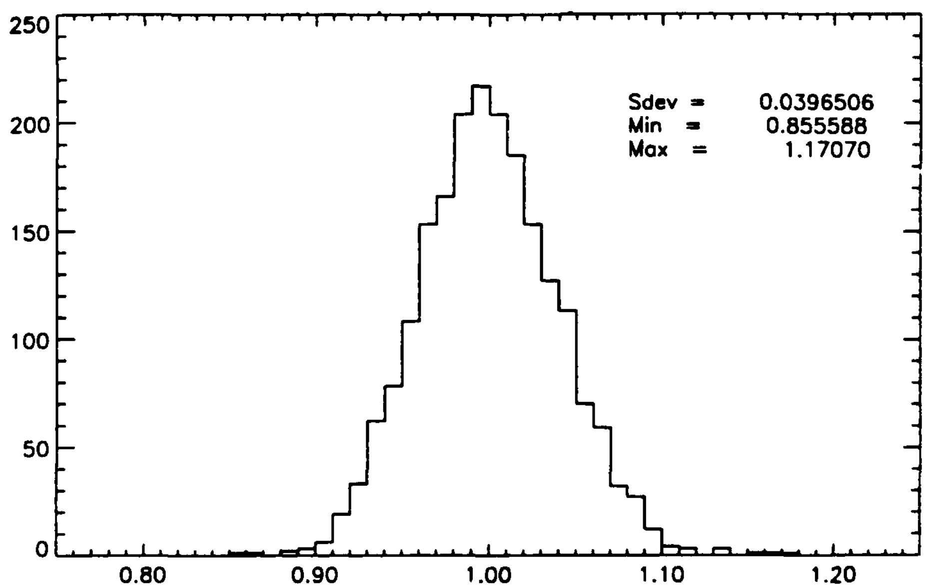


Figure 4: Intensity distribution for a 64x32 spot array measured by CCD system.

Packaged optical interconnection system based on photorefractive correlation

Hidemi Takahashi*, David Zaleta, Jian Ma, Joseph E. Ford, Yeshaiahu Fainman, and Sing H. Lee
University of California, San Diego
Department of Electrical and Computer Engineering, La Jolla, CA 92093
Phone:619-534-2325

1. Introduction

An opto-electronic multiprocessor system consisting of a large number of electronic processing elements with optical inputs and outputs is promising for future parallel processing because this architecture removes the limitations of the VLSI technology by providing the interconnects in three dimensions.[1] For this purpose, a free-space optical interconnect system based on photorefractive correlation has been proposed called the correlation matrix-tensor multiplier(CMTM) algorithm.[2] To realize a practical system, ease of alignment, reliability of the optical system, and ruggedness are required. One method to satisfy these requirements is a planar integration of the optical components on a glass substrate.[3] This paper describes the demonstration of a packaged free-space optical interconnection system based on the CMTM algorithm as one step towards a future realization of practical and rugged system.

2. CMTM System

In the CMTM algorithm, the two dimensional input array is correlated with an interconnection control pattern to generate the output array. The control pattern is designed to generate the desired interconnection. A random phase code is used for both the input and the control pattern to suppress the correlated output at the undesired positions. The correlation is obtained by the photorefractive four-wave mixing. Figure 1 shows the schematic of the CMTM optics. The interference pattern between the plane wave and the Fourier transform of the control pattern is recorded in a $LiNbO_3$ crystal. When the Fourier transformed input illuminates the crystal, the diffracted output is the correlation between the input and the control pattern. The optical components for the CMTM are therefore a control pattern, a phase code, a Fourier transform lens, and a $LiNbO_3$ crystal. The control pattern is only necessary for recording. In the interconnection mode, only one beam , the input beam, is necessary. Reconfiguration between pre-stored interconnections is possible by using the wavelength multiplexed storage.

Because the CMTM is suitable for a free space interconnection between two chips, this system is packaged on a glass substrate in a practical realization. We have constructed an experimental packaged CMTM system as shown in Fig.2. A phase code, two reflective CGH Fourier transform lenses, a $LiNbO_3$ crystal, and a CCD camera are attached to one surface of a glass substrate. The input beam is Fourier transformed into the $LiNbO_3$ crystal, and the diffracted output from the crystal is Fourier transformed to generate the correlation between the input and the control pattern, which is the interconnected output pattern.

3. Optical Components for Experimental System

A. Phase code and control pattern

The phase code and the control pattern are transmission plates fabricated by the electron-beam lithography and reactive ion beam etching. The input array is 5×5, therefore, the control pattern

*On leave from Matsushita Research Institute, Tokyo Inc., 3-10-1 Higashimita, Tama-ku, Kawasaki 214, Japan, Phone:+81-44-911-6351

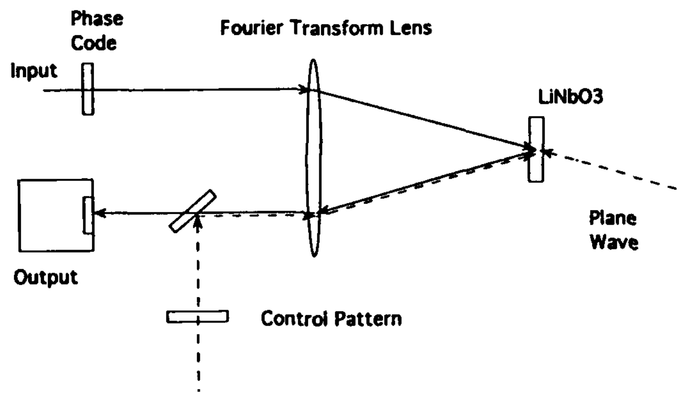


Fig.1. Schematic of the CMTM optics.

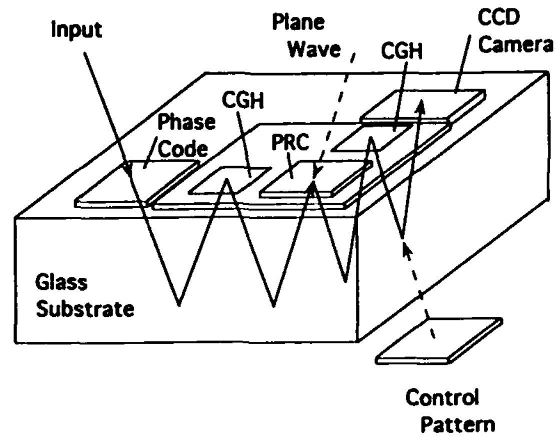


Fig.2. Schematic of the experimental packaged CMTM system.

has 25×25 pixels. The size of one pixel is 250 μm×250 μm. Each pixel consists of 5×5 sub pixels of the phase code.

B. $LiNbO_3$ crystal

The crystal used in the experimental system is 0.05 % Fe doped $LiNbO_3$ z-cut crystal, and 20×20×2 mm in size. In the CMTM system, reflection holograms are recorded in the $LiNbO_3$ crystal. The diffraction efficiency of the crystal for the reflection hologram is measured. The result is shown in Fig.3. The light source is Argon laser operated at 514.5 nm. The maximum efficiency obtained was 23 %.

The solid line in Fig.3 shows the theoretical fitting calculated from Kogelnik's coupled wave theory.[4] The diffraction efficiency η is given as a function of the recording time t by

$$\eta = \frac{1}{[\xi / v + \sqrt{1+(\xi / v)^2}\coth\sqrt{v^2+\xi^2}\,]^2} \quad , \tag{1}$$

$$v = \pi d \Delta n_s (1 - e^{-t/\tau}) / \lambda \cos\theta \quad , \tag{2}$$

$$\xi = \alpha d / \cos\theta \quad , \tag{3}$$

where d is the thickness of the crystal, Δn_S the saturation value of the refractive index modulation, τ the recording time constant, θ the incident angle of the beam, α the absorption coefficience. According to this theoretical fitting, Δn_S is derived to be 1.1×10^{-4}.

C. CGH Fourier transform lenses

The CGH Fourier transform lenses were designed from Code V. In general, Code V enables the design of a phase function

$$\phi = \frac{2\pi}{\lambda} \sum_i \sum_j C_{ij} x^i y^j \quad . \tag{4}$$

We utilized the orthogonal cylindrical diffractive lens design with only C_{20} and C_{02} terms (x^2 and y^2 coefficients)[5], so the phase function was given by

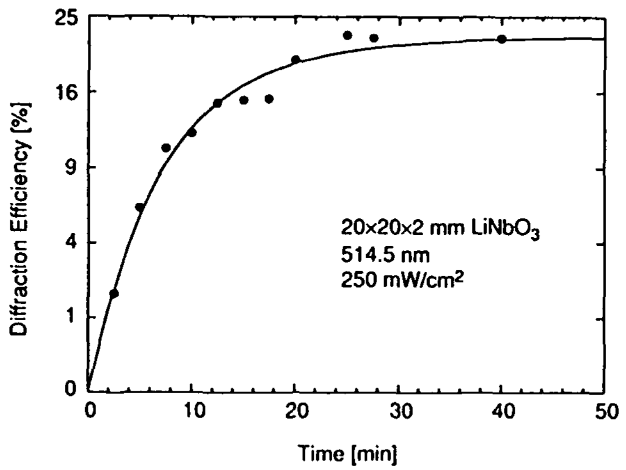


Fig.3. Measured diffraction efficiency as a function of the recording time for a reflection hologram in the LiNbO3 crystal.

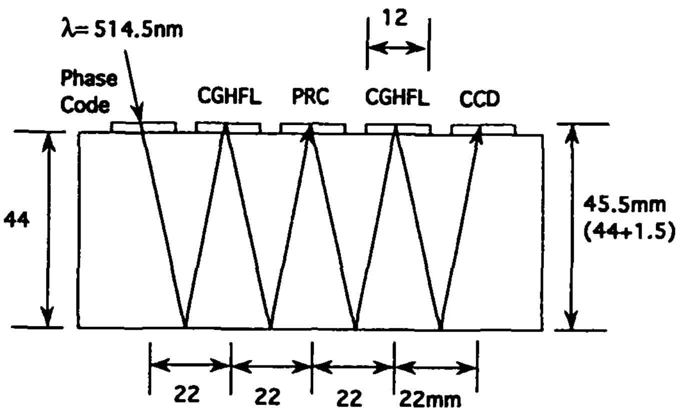


Fig.4. Design parameters of the CGH Fourier transform lenses. CGHFL:CGH Fourier transform lens, PRC:photorefractive crystal.

$$\phi = -\frac{\pi}{\lambda}\left(\frac{x^2}{f_x}+\frac{y^2}{f_y}\right) , \qquad (5)$$

where f_x and f_y are the focal length in x and y directions. To minimize the spot diameter at the focal point for off-axis rays, f_x and f_y are determined by an optical design program.

Figure 4 shows the design parameters of the CGH. For simplicity, the optical path is shown to be straight instead of 90 ° rotation by the crystal. The thickness of the glass substrate is 44 mm, and the CGHs are designed on the top surface of a 1.5 mm thick glass plate. The spacing between each component is 22 mm. The aperture of CGHs is 12×12 mm. The phase levels of the CGHs are binary, and the measured diffraction efficiency for the 1st order was 32 %.

4. Experimental Results and Discussions

The experimental packaged system is constructed of which photograph is shown in Fig.5. The dimension of the glass substrate is 80×80×44 mm. To estimate the uniformity of the output signals for all the output points, all to all interconnection between 5×5 arrays are examined. There is noise at the center of the outputs. This is the 0th order diffraction by the CGH. Because the CGH has only two phase levels, the 0th order diffraction is not small. This will be eliminated by using multi-phase-levels CGHs.

Fig.5. Photograph of the experimental packaged CMTM system.

Another way of improvement is to use an off-axis

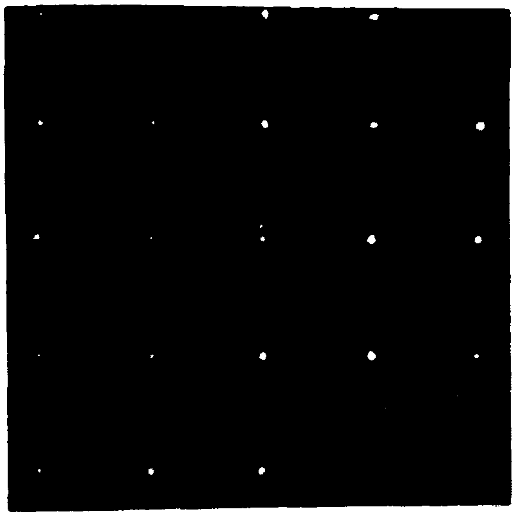


Fig.6. Result of all to all interconnection between 5×5 arrays.

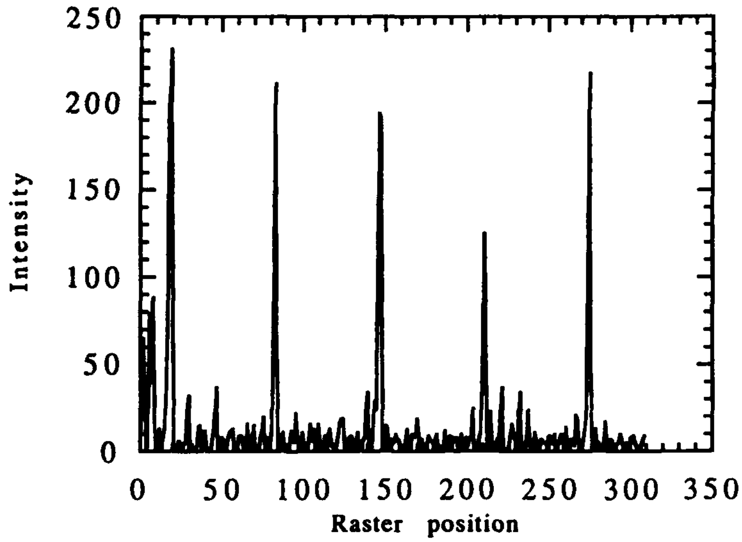


Fig.7. One horizontal scan output from the CCD camera.

configuration instead of the on-axis configuration to separate the 0th order diffraction from the 1st order spatially. The result of all to all interconnection between 5×5 arrays by the off-axis configuration is shown in Fig.6. The noise at the center is eliminated, but the intensity of all the points are not uniform. Because the CGH is designed for the on-axis configuration, aberrations are large when it is used in the off-axis configuration. The utilization of off-axis CGH lenses designed to have smaller aberrations is the way for further improvements. An example of one horizontal line scan output of the CCD camera is shown in Fig.7. The scanning position is the center row of the five rows of output points. Except the noise on the left(0th order diffraction beam), the maximum value of SNR is 10. The SNR of 20 is obtained for each output point if the alignment is optimized for that point. The SNR obtained by experiments using conventional optics is 100. The difference between these SNR values are mainly due to the other order diffraction beams by the CGH.

5. Conclusions

We have demonstrated the packaged free-space optical interconnection system on a glass substrate based on the CMTM algorithm. The experiments with this system show the possibility of the packaging of optoelectronic systems in this scheme. For the realization of a practical optoelectronic multiprocessor system, further work in this approach is essential.

References

[1] F.Kiamilev, S.C.Esener, R.Paturi, Y.Fainman, P.Mercier, C.C.Guest, and S.H.Lee, Opt.Eng. **28**, 396 (1989).
[2] J.E.Ford, Y.Fainman, and S. H. Lee, Optics Letters **15**, 1088 (1990).
[3] J.Jahn and A.Huang, Appl.Opt. **28**, 1602 (1989).
[4] Herwig Kogelnik, Bell.Syst.Tech.J. **48**, 2909 (1969).
[5] K.Urquhart and S. H. Lee, Submitted to Appl.Opt. (1992).

Polarization-Selective Computer-Generated Holograms for Optical Multistage Interconnection Networks

Joseph Ford, Fang Xu, Ashok Krishnamoorthy, Kristopher Urquhart and Yeshayahu Fainman
University of California San Diego, ECE Dept., La Jolla CA 92093-0407; (619) 534-8909

1. Introduction

Computer-generated holograms (CGH) fabricated as phase-only optical elements have proven to be useful in implementing free space optical interconnections for parallel computing and communications. Such elements are normally capable of implementing a single fixed optical interconnect. The performance of such systems could be enhanced with switching optical interconnects that do not require signal detection and regeneration. A polarization-selective CGH with independent impulse responses for the two orthogonal linear polarizations would introduce an additional degree of freedom useful in designing such systems. An array of such polarization-selective CGH could be combined with electro-optic polarization rotators to build an electrically-controlled optical switch array for implementing a passive multistage interconnection network (MIN) supporting fast data transmission bandwidths.

Optically recorded polarization-selective holographic optical elements have been obtained using various media. However, CGH offer possible advantages over optically recorded elements in efficiency, repeatability, and generality. Our goal is a polarization selective CGH. Our physical approach is related to that of Ohba et al,[1] where proton exchange in lithium niobate was used to create a birefringent grating compensated by a surface dielectric layer. This approach produces a single hologram, invisible to the orthogonal polarization. In contrast, our approach[2] yields CGH with arbitrary functionality for each of the two orthogonal linear polarizations.

2. Design and fabrication of the birefringent CGH

A conventional phase CGH can be fabricated by etching an isotropic substrate with a surface relief pattern which will impose the desired phase delay on an incident optical wavefront. To build an optical element with arbitrary function for each of the two orthogonal polarizations, we need to apply an independent phase delay for each polarization at each pixel. This additional information must be physically encoded into the substrate. We use two birefringent substrates, with their etched surfaces in contact, to hold this information (see Figure 1).

These two substrates can apply an arbitrary phase for the two orthogonal linear polarizations of a transmitted light ray. Consider the case of one birefringent and one isotropic substrate, where the polarization of the incident light is either aligned with or perpendicular to the birefringent substrate's optical axis. A ray transmitted through the birefringent substrate will have a different phase delay for each polarization because the index of refraction is different. The phase angle between the two polarizations and the absolute phase delay of the rays depends on the thickness, and hence the etch depth, of the birefringent substrate. This etch depth is chosen to obtain the desired final phase angle between the two polarizations. The ray then passes through the isotropic substrate, where light of either polarization is delayed by the same phase angle, again depending on the etch depth. This etch depth is chosen to bring one polarization to the desired phase angle. Since the relative delay between polarizations is unaffected by the isotropic substrate, the phase angle of the orthogonal polarization is simultaneously brought to the final desired value.

The etch depths are calculated for the more general case of two birefringent substrates as follows. At each pixel, the optical path difference relative to the unetched region is given by $\Phi_o = d_1(n_g-n_o) + d_2(n_g-n_o')$ for ordinary polarization and $\Phi_e = d_1(n_g-n_e) + d_2(n_g-n_e')$ for extraordinary polarization, where the etch depths are d_1 and d_2 for the first and second substrates, the indexes of the first substrate are n_o and n_e, the indexes of the second substrate are n_o' and n_e', and the gap material is assumed to be isotropic with index n_g (see Figure 1). Solving these two linear equations in two unknowns yields the required etch depths for any combination of phases:

$$d_1 = \frac{(n_o'-n_g)\Phi_o - (n_e'-n_g)\Phi_e}{(n_e-n_g)(n_o'-n_g) - (n_e'-n_g)(n_o-n_g)} \qquad d_2 = \frac{(n_o-n_g)\Phi_o - (n_e-n_g)\Phi_e}{(n_e'-n_g)(n_o-n_g) - (n_e-n_g)(n_o'-n_g)} \tag{1}$$

The positive and negative values produced by these equations can be normalized to raise the highest feature to zero depth. We designed a binary phase hologram fabricated in lithium niobate ($n_o = 2.33$ and $n_e = 2.24$ at $\lambda = 514.5$ nm), using identical perpendicularly-oriented birefringent substrates. Such an arrangement reduces the etch depths (in this case, from 12 µm to 3 µm) and allows the two substrates to be etched simultaneously. There are four distinct etch values for the hologram (0, 1.37, 1.47, and 2.84) but they can be achieved with only two etches, because the fourth value is the sum of the second and the third. In general, with symmetric substrates the number of etches required to fabricate an N-level polarization-selective hologram is $2\log_2 N$, exactly twice the minimum number of etches with a simple N-level hologram for a single polarization.[3]

3. Demonstration of the 1x2 switch for multistage interconnection networks

These BCGH elements can be used to construct a self-routing 2x2 binary switch for free space optical interconnections. The switch consists of two dual polarization selective holograms separated by a polarization rotator. Two input beams are combined and possibly focused into the polarization rotator, then separated and redirected by the second dual hologram. The polarization rotator either exchanges the two beams, or not, depending on the applied voltage. A network constructed of such switches would be circuit-switched, since optical information is routed without detection and signal regeneration. The data modulation rate is limited only by the input sources and output detectors, not by the network itself. However, the switches need to be efficient to preserve signal intensity.

We designed a dual beamsplitter to deflect one polarization vertically and the other horizontally. The hologram was 5mm x 5mm, with a design wavelength $\lambda = 514.5$ nm and a minimum feature size of 10 µm. Kinoform patterns were calculated and converted to binary phase for each polarization. Then the two kinoforms were combined to produce mask patterns for two etches of 1.37 and 1.47 µm. Chrome masks were generated by electron-beam lithography and transferred to a photoresist on the substrates by contact print lithography. The patterns were etched into the lithium niobate substrate using ion beam milling. The accuracy of the etch depth was approximately +/- 1%. Some lateral etching occurred, producing fringes with sloping sidewalls and somewhat reduced width (2 µm per side). The substrates were aligned in a standard mask aligner to within 2 µm, then permanently joined with UV curing epoxy. The maximum first order diffraction efficiency (not including surface reflections) and the polarization contrast ratio of the dual beamsplitter were measured to be 6% and 40:1 respectively. Less than 0.2% of the intensity was lost into the zeroth order.

This element was used in the 1x2 optical switch configuration shown in Figure 2. A beam chopper modulated a laser beam to emulate a data signal. This signal was then transmitted through a liquid crystal polarization rotator, which switched the signal polarization at a lower frequency representing the network reconfiguration rate. Upon transmission through the BCGH, depending on the signal polarization, the modulated laser beam was switched between two output photodetectors. The output of the two detectors is shown in the inset photograph. The contrast ratio of the liquid crystal modulator was 200:1, of the BCGH was 40:1, and of the full switch was 25:1. The efficiency of this switch was limited to 3% due to fabrication errors, but the theoretical hologram efficiency approaches 100%. Further work on improved fabrication techniques is under way.

4. Multistage interconnection network architecture

We have studied different optoelectronics MIN architectures and control algorithms[4] for constructing scalable MINs with the BCGH switch technological constraints. Two information routing algorithms are possible: centralized control, where an external control processor determines and configures all switch states, and distributed control with self-routing packet headers. In the latter case, the switch settings are calculated external to the network but the process of setting the switches is implemented using packet headers which propagate through the network and configure the "smart" opto-electronic spatial light modulator switch arrays during a dedicated header time interval. Once the switches in the MIN are set, the information may be both sent to and retrieved from the output nodes through the same optical system at the speed of light, with a data modulation rate which is not limited by the network.

We have evaluated different interconnection network architectures (i.e., crossbar, N-stage planar, Batcher-Banyan, Benes, and Dilated Benes) in conjunction with their performance using BCGH technology in a passive all-optical network. For this study we computed the network attenuation losses and the network SNR vs the total number of switches using single switch insertion loss of -0.638 db and switch contrast ratio of 20 db. If we assume the maximum acceptable attenuation to be 30dB and the minimum SNR to be 11dB [5] then the maximum scalability (dotted lines) of the architectures can be estimated as shown in Figures 3 and 4.

We concluded from this study that the dilated Benes network[6], which is a rearrangeably non-blocking MIN, is well suited for large scale implementations of circuit-switched photonic interconnection networks using BCGH switching elements. Most importantly, the size of the network realistically possible using BCGH technology may be as large as 16,384 inputs with a 16 GHz data rate, well above the size and data modulation rate achievable with all-electronic networks.

5. Conclusions

In conclusion, we have introduced a method for constructing polarization-selective CGH using surface-relief etching of birefringent substrates. We tested this approach for a dual binary phase CGH fabricated in lithium niobate, yielding a diffraction efficiency was 6% and a 40:1 contrast ratio between polarizations. This element was used to experimentally evaluate a self-routing 1x2 polarization switch. Finally, we have examined technological implications of BCGH on optical MIN architectures.

This work was funded by Rome Laboratory under Grant F-30602-91-C-0094.

1. A. Ohba, Y. Kimura, S. Sugama, R. Katayama, and Y. Ono, Appl. Opt. 29, 5131-5135, 1990.
2. J. Ford, F. Xu, K. Urquhart, and Y. Fainman, *"Polarization selective computer generated holograms,"* accepted by Optics Letters September 1992.
3. G. Swanson, M.I.T. Lincoln Laboratory Tech. Rep. 854, August 1989.
4. T. Cloonan, F. McCormick, and A. Lentine, OSA Tech. Dig., Topical Meeting on Photonic Switching, Salt Lake City, 1991.
5. N. Ailawadi, in Frontiers in Computing Systems Research Vol. 1, S. Tewksbury Ed., Plenum Press 1990.
6. K. Padmanaphan and A. Netraveli, IEEE Trans. on Comm. C-30(12), December 1987.

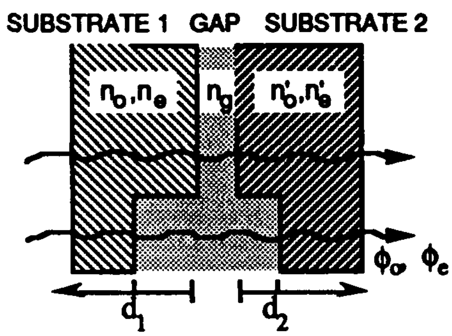


Figure 1. BCGH construction.

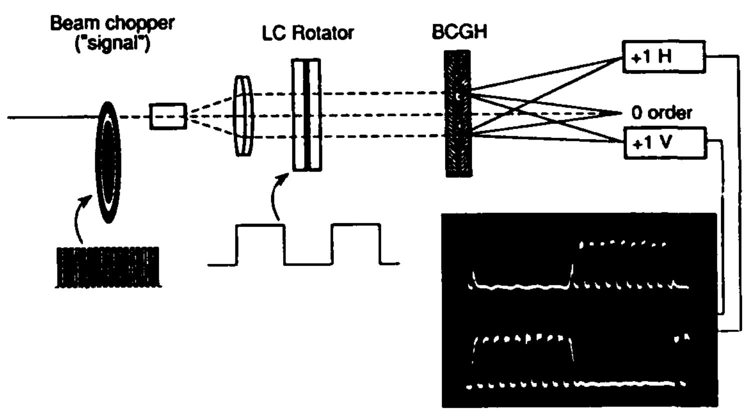


Figure 2. 1x2 polarization switch demonstration.

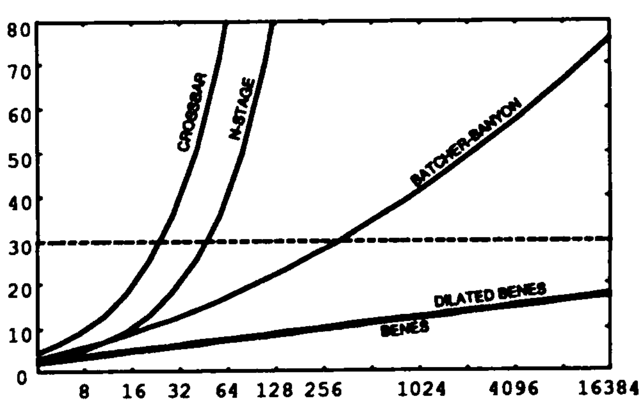


Figure 3. Network attenuation in db vs network size.

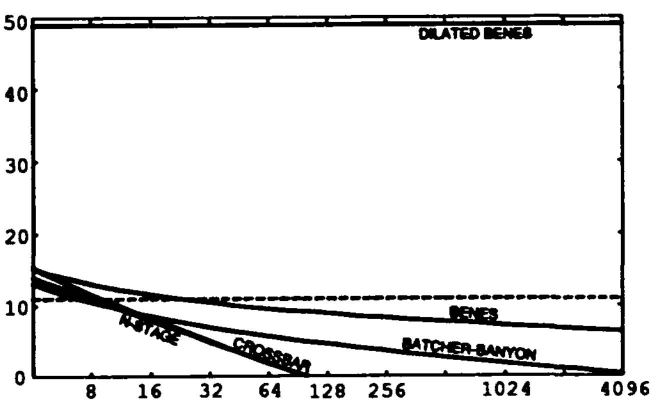


Figure 4. Network SNR in db vs size.

Parallel Systems: 2

OThD 4:00pm-6:10pm
Grand Ballroom West

Jacek Chrostowski, *Presider*
National Research Council, Canada

Optically Interconnected Multichip Modules using Computer Generated Holograms

Michael R. Feldman, James E. Morris, John Childers, Mouna Nakkar and Fouad Kiamilev
University of North Carolina at Charlotte, EE Dept., Charlotte, NC 28223

1. Introduction

Current electrical multichip modules will limit the performance of many next-generation highly parallel computational systems. Since such systems require highly dense connection networks containing many long distance connections, the minimization of the area, power and time delay of the chip-to-chip and module-to-module interconnects is of utmost importance. By replacing particular electrical intramodule and module-to-module connections with optical communication links, this communication bottleneck can be relieved. Optical interconnects have the potential to increase communication speed, and reduce the volume, crosstalk and power dissipation of the connections [1-5].

A multichip module (MCM) designed specifically to meet the demands of high performance processor array systems has been developed. The system consists of Computer Generated Holograms (CGH's), GaAs laser array chips, and detectors integrated onto silicon Integrated Circuits (IC's). All components are incorporated into a package (similar to an existing multichip module design) with a water-cooled heat sink. All intramodule chip-to-chip connections longer than a particular line length and all intermodule connections are implemented optically.

2. Concept

The structural design of a multichip module with holographic optical interconnects has been developed. The design is illustrated in Fig. 1[6-8]. A set of silicon VLSI chips are attached to the bottom side of the substrate in Fig. 1 with flip-chip (C4) technology. Small (~300 µm x 1 cm) GaAs laser array chips are also attached to the bottom side of the substrate through flip-chip soldering. One GaAs laser chip is placed next to each silicon IC. A layout of the IC's on the bottom side of the substrate is shown in Fig. 2 [7,8]. The GaAs chips contain arrays of laser diodes. The center-to-center spacing between adjacent lasers is 300 µm in order to limit the power dissipation per unit area and to provide sufficient area to place a contact pad adjacent to each laser. The silicon IC's contain integrated photodetectors.

The bottom side of the substrate contains several layers of metal lines for power and signal distribution as well as flip-chip bump pads. A planar heat sink is used to cool both the silicon and GaAs IC's as in conventional flip-chip MCM's.

The top side of the substrate contains a Computer Generated Hologram (CGH). The CGH is divided into subholograms. One subhologram is placed over each laser ("laser subhologram") and one subhologram is placed over each detector ("detector subhologram"). The subholograms function to change the shape of the incident optical signal beam. Each laser illuminates the subhologram located directly above it. The laser subhologram divides the light into F beams (to provide a fanout of F), and directs each beam onto the appropriate detector subhologram after reflection off of the planar mirror located approximately 2 cm above the substrate (Fig. 1). Each detector subhologram acts as a single lens and focuses the incident beam onto the detector located directly below it. This arrangement is termed a Double Pass optical interconnect system after Refs. 9,10. (Note that this system is similar to a substrate mode interconnect system[11]). Optical connections can be made between chips within a given module or between chips located on different modules.

The interconnection density of a double pass holographic system is calculated in Ref. 9 as a function of CGH deflection angle and number of transmitters and detectors. Equations derived in Ref. 9 indicate that for a 50 degree CGH deflection angle, the double pass system can provide over 80,000 connections (assuming an average fanout of 4) with a 10 cm diameter CGH.

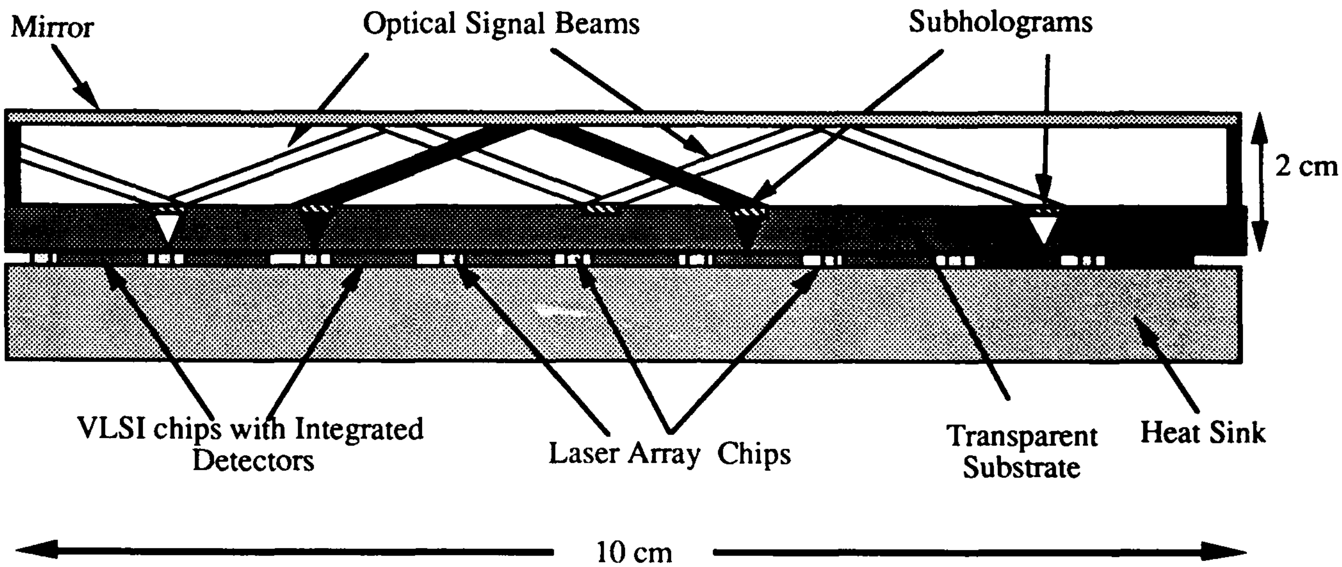


Fig. 1. Cross-sectional view of an optically interconnected MCM.

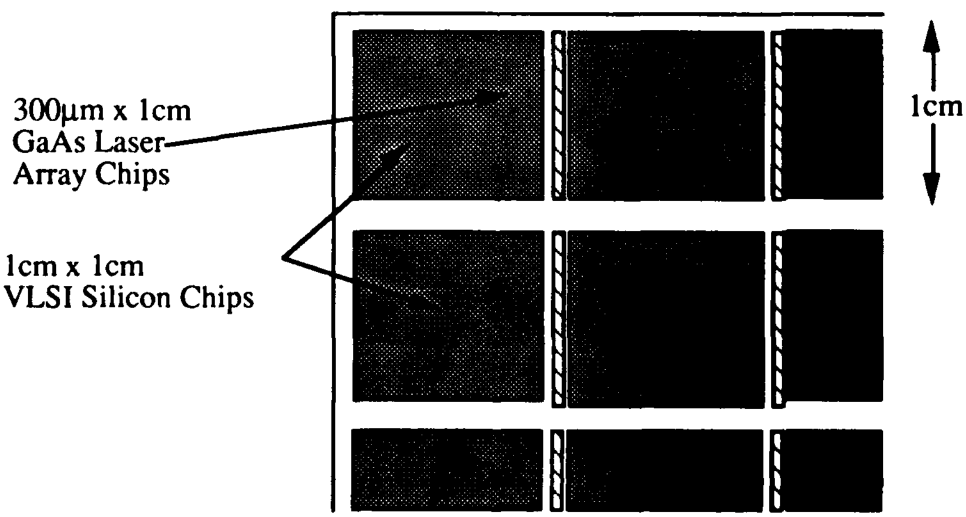


Fig. 2. Layout of Chips on MCM substrate.

3. Prototype Development

Development of a fully functional prototype optically interconnected MCM is currently underway at UNC-Charlotte and MCNC. The initial prototype employs edge-emitting laser arrays that are flip-chip bonded to the substrate along with micro-mirrors to redirect the light output. Photodetectors are integrated onto 1.25 μm CMOS chips that are also flip-chip bonded to the substrate. Alignment of the components to the substrate is achieved by flip-chip bonding techniques.

The prototype module will have 54 optical links, made up 33 1:1 connections of various lengths and angles and 21 connections with fanout of 2 to 3. These various connections will enable the experimental determination of cross-talk, signal-to-noise ratio, alignment sensitivity and efficiency of the optical links.

New CGH encoding methods have been developed to provide high diffraction efficiency and tightly focussed beams over a range of laser diode emission wavelengths. One of these methods, known as Radially Symmetric Iterative Discrete On-axis encoding (RSIDO), was used to design an F#/1 CGH with 89% diffraction efficiency while operating over a ±10nm bandwidth[12]. Another encoding method termed SRP (Segmented Radial Partitions) was developed to allow application of the RSIDO encoding method to non-radially symmetric elements. Experimental results indicate that, for example, for a CGH deflection angle of 27 degrees, the diffraction efficiency can be increased to ~80% with the SRP method compared to ~16% with conventional encoding methods.New photodetector designs and fabrication methods have been developed to improve the performance of photodetectors integrated with CMOS VLSI chips. Previous detector gates integrated onto CMOS chips have demonstrated response times (rise and fall) times greater than 30 - 40 nsec [13]. We have experimentally measured response times (both rise and fall) of under 5 nsec. Modulation at datarates of 180 Mbits/sec have been demonstrated[14].

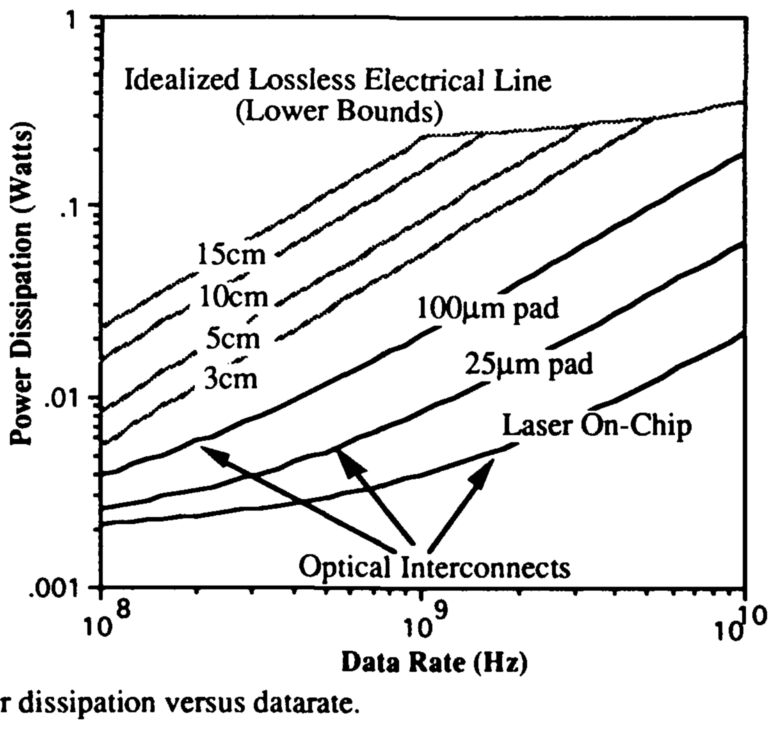


Figure 3. Power dissipation versus datarate.

4. COMPARISON WITH ELECTRICAL INTERCONNECTS

In Ref 2, the power dissipation and switching energy of optical interconnects was compared to that of electrical interconnects for connections within a single IC. This comparison has been extended to compare the power dissipation of optical and electrical interconnects for connections between chips within an MCM [7,8]. Results are summarized in Fig. 3. The optical interconnect curves take into account the power dissipation in the laser diode, in the detector circuitry and in the charging of the bonding pads and metalized communication links between the silicon laser drive and the GaAs laser. The electrical interconnect curves were obtained from 2 different models. When the signal wavelength is longer than twice the line length, a lumped RC model was employed. When the wavelength is shorter then twice the line length an ideal lossless perfectly matched transmission line model is employed. Thus, the electrical interconnect curves represent lower bounds on the power dissipation for electrical connections. From Fig. 3, it is evident that our current optically interconnected MCM approach (with pads of 100 μm side length) can provide an order of magnitude decrease in the power dissipation compared to that of electrical interconnects, with line lengths longer than 5-6 cm, for data rates up to ~3-4 GHz. For smaller bonding pads, the one order of magnitude decrease in power dissipation (compared to that of electrical interconnect lower bounds) can be maintained for data rates in excess of 10 GHz. It can also be seen that integration of optoelectronic transmitters onto the chip has only a small impact (less than a factor of 2-3) until datarates in excess of approximately 5-10 GHz are employed.

References

1. J. W. Goodman, F. I. Leonberger, S. Y. Kung and R. A. Athale, "Optical Interconnections for VLSI Systems," Proc. IEEE, vol. 72, pp. 850-866, 1984.

2. M. R. Feldman, S. C. Esener, C. C. Guest and S. H. Lee, "A Comparison Between Optical and Electrical Interconnects Based on Power and Speed Considerations," Appl. Opt., vol. 27, pp. 1742-1751, May 1,1988.

3. M. R. Feldman, C. C. Guest, T. J. Drabik and S. C. Esener, "A Comparison Between Optical and Electrical Interconnects for Fine Grain Processor Arrays based on Interconnect Density Capabilities", Appl. Opt., vol. 28, pp. 3820-3829, 1989.

4. L. A. Bergman, W. H. Wu, A. R. Johnston, R. Nixon, S. C. Esener, C. C. Guest, P. Yu, T. J. Drabik, M. R. Feldman and S. H. Lee, "Holographic Optical Interconnects for VLSI," Opt. Eng., vol. 25, pp. 1109-1118, 1986.

5. D. Z. Tsang, "One-Gigabit per second Free-Space optical Interconnection," Appl. Optics, vol. 29, pp. 2034-2037, 1990.

6. Michael R. Feldman, "Diffractive Optical Elements for Optical Interconnect Applications," SPIE conference on Miniature and Micro-Optics and Micromechanics, San Diego, CA, 1993.

7. M. R. Feldman, I. Turlik, G. Adema, J. E. Morris, N. Hangzo and W. H. Welch, "Holographic Optical Interconnects for VLSI Multichip Modules," in 42nd Electronic Components & Technology Conference 1992, Vol. 1, 513 - 518, 1992.

8. M. R. Feldman, I. Turlik, G. Adema, J. E. Morris, and W. H. Welch, "Holographic Optical Interconnects for VLSI Multichip Modules," accepted for publication in IEEE Transactions on CHMT 1992.

9. M. R. Feldman and C. C. Guest, "Interconnect Density Capabilities of Computer Generated Holograms for Optical Interconnection of Very Large Scale Integrated Circuits", Appl. Opt., vol. 28, pp. 3134-3137, 1989.

10. M. R. Feldman and C. C. Guest, "Design of Computer Generated Holograms for a Shared Memory Network," Proc. of SPIE, vol. 1052, 1989.

11. R. K. Kostuk, Y.-T. Huang, D. Heterington and M. Kato, "Reducing Alignment and Chromatic Sensitivity of Holographic Optical Interconnects with Substrate-Mode Holograms," Appl. Opt.,vol. 28, p. 4939, 1989.

12. W. Hudson Welch, J. E. Morris, and M. R. Feldman, "Iterative Discrete On-axis Encoding of Radially Symmetric Holograms," submitted to JOSA-A.

13. W. Wu, L Bergman, A. Johnston, C. Guest, S. Esener, P. Yu, M. Feldman, and S. Lee, "Implementation of Optical Interconnections for VLSI," IEEE Transactions on Electron Devices, Vol. ED-34, No. 3, March 1987.

14. J. Childers, James E. Morris, Michael R. Feldman "CMOS compatible photodetectors for optical interconnects," SPIE O/E LASE 93, Los Angeles, January 1993.

MASSIVELY PARALLEL COMPUTERS USING HOLOGRAPHIC MATRIX AS A PHOTONIC BACKPLANE

M. Charrier, B. Houssay, T. Lemoine, S. Paineau - Thomson CSF - RCM Division
178, avenue Gabriel Peri - 92242 Malakoff - France - Tel (33-1) 49-65-30-30

0 - Massively parallel computers in airborne radars : A communication problem

The computing power presently required in building airborne radars computers exceeds 1GOps in 10 liters.

This is mainly justified for signal processing and the introduction of the electronic beam forming in the next generation of systems should increase it up to 100 GOps in the same volume.

To reach this objective, the designers will use microprocessors which are more and more powerful, so that the quantity of processors won't increase in the same proportions that the calculation power - even if it will go over 100 processors. This requires that inter-processors communication problems are solved so that each processor is used to optimum capacity: the problem arising in the future generations of computers will be - and already is - a problem of communication (2).

To interconnect these processors, the optical technologies offer possibilities which are compatible with new computer's architectures using large quantities of processors: The optical interconnect is well adapted with a broadcast distribution of signals, which have been the subject of an ESPRIT II program exposed in this document. Furthermore, optics is of great interest for clock distribution with reduced skew between receivers, especially in synchronous machines.

The conventional "electrical" technologies like MCMs are perfectly suitable for needs in connections density at the surface of a board. Frequency is the only limitation, related to impedance mismatches in the case of signal division. So, one specific interest of optical interconnect on a board is the distribution of a high frequency clock signal in a synchronous machine.

In a backplane, the problem is definitely more serious. To the problem of the 1-->N signals distribution - which still subsists (clock distribution, broadcast communication networks) is added the casual impedance mismatch for high frequency signals related to the interconnect technology used ; it can only be solved by the loss of considerable space : a coaxial connection consumes ten times as much space as a conventional connection.

Optical interconnect offers an alternative which seems to be unavoidable. The propagation of an optical signal - in multimode or free-space operation - is, aside from losses, practically insensitive to impedance mismatch.

1 - Holographic-Type Photonic Backplane

The technology presented here concerns the realisation of a photonic motherboard as developped by THOMSON CSF with the support of the French (DRET) and European (ESPRIT II - OLIVES) administrations. This motherboard is composed of a glass plate perpendicular to the boards carrying the processors. In the motherboard, the light beams are propagating in a half guided mode by total internal reflections. CGH (Computer Generated Holograms), located right on the glass vertically above the daughterboards, are making up the optical input-outputs of the glass plate(fig. 1); the glass plate is then known as a "holographic matrix".

In the simplest case, a diffraction pattern reflects a perpendicular incident beam with a certain angle (emission or type-E hologram) or pick up perpendiculary a part R(1) of an oblique beam (reception or type-R hologram) (fig. 1). If more complex operations are involved, for example "deviate a beam",calculating the pattern and, especially, producing it become more complicated (type D hologram).

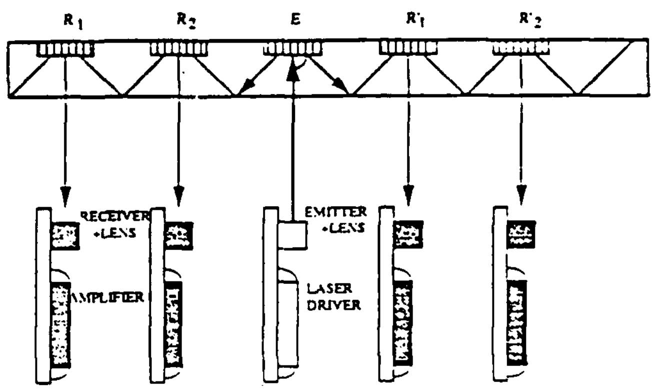


fig -1-

2. Producing a Holographic Matrix : A Practical Description

A holographic matrix is a plate of glass, one face of which is etched with localized diffractive elements. The material have to be transparent at the wavelength used (typically 1.3 µm). Borosilicate or quartz offers a good compromise (index = 1,5)(1).

The glass plate must not be too thin in order not to increase the quantity of reflections which inevitably produces losses (1 mm is also a good compromise).In addition, a plate that is too thin becomes delicate to process when creating holograms.

The holographic elements are initially chromium patterns implanted at the surfaceof the glass plate. The fineness of the lines making up the diffractive array is directly related to the wavelength of the light used and the function of the hologram (type E, R or D). For a wavelength of 1.3 µm, the creation of type E or R interfaces requires a line pitch of 1.2 µm. For a type D holograms, the pitch can locally go down to 0,85µm, which is the best that industrial E-beam processes can do. The use of short wavelengths around 0.8 µm is foreseen in the next few years to get advantage of the low cost of CD lasers.

The glass plate with the required chromium patterns (which is nothing more than a semiconductor mask) is then etched using a dry process which is more easily controlled for depths on the order of a few hundred nanometers (3).

The depth is the determining factor that controls the amount of energy diffracted in the various directions . Requirements call for deeply-etched holograms (more than 300 nm) for the emission patterns, but reception holograms must be only slightly etched (less than 100 nm). The necessity to have simultaneously at least these two types of holograms complicates the manufacturing process by increasing the quantity of operations. The process remains nonetheless fully industrial as it has been extensively used by the semiconductor industry. The production process is carried out in batches as opposed to an optical fiber backplane which must be produced one at a time. In addition, and as opposed to silicon circuits produced by the same kind of processes, holograms are not easily affected for defects such as conductor discontinuity or bridging, considering that they are microscopic: the process is therefore not affected by problems of yield which limit the semiconductor industry and is thus compatible with the industrial manufacture of large-size plates (several dm2).

3 - ESPRIT II "OLIVES" program : Broadcast 9 --> 9 boards network.

Several prototypes have been built with this technology. One of the most interesting of these is a breadboard of an optical backplane for nine daughterboards providing a broadcast-type communication network (fig.2). Each daughterboard supports an emitter distributing an optical signal to the other eight boards simultaneously. Each board therefore supports also eight receivers. Provided with the required low-level protocol, each board can accept or refuse the received messages. This kind of function is particularly interesting for future computers architectures and, in particular, for data processing . By optimizing the routing schemes, it allows messages transmission between each processors at a high rate with low latency time.

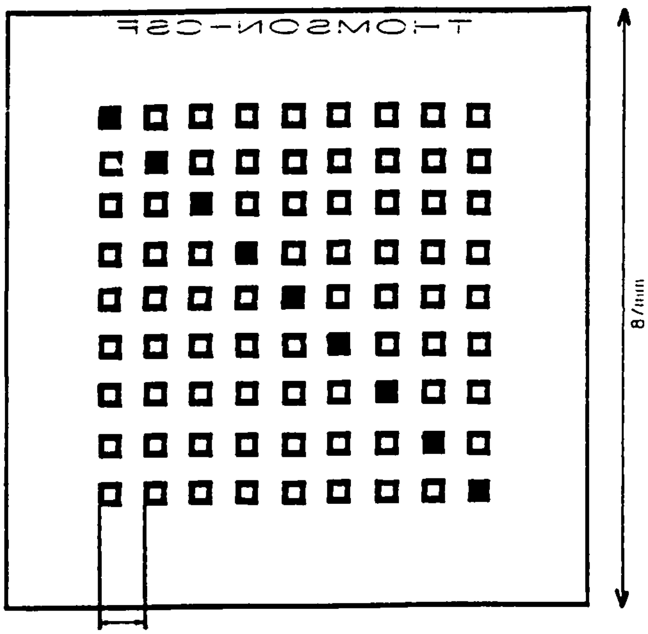
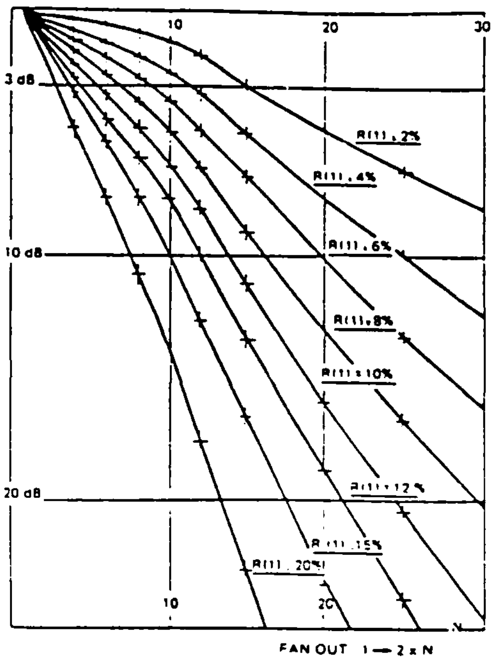


fig -2-

Optical backplane



Input energie on the first receiver (N=1)

R(1) %	2	4	6	8	10	12	15	20
P1 µW	70	140	210	280	350	420	525	700

[Output energie of the laser diode : 8 mW]

fig -3-

Dynamic between the 1^{st} and the N^{th} receivers, for various diffraction efficiencies R(1).
P1 is the incident power on the first CGH.
(pitch between cards: 6,35mm)

The breadboard is composed of nine ceramic MCM in a rack with a pitch of 6.35 mm.
Each board is provided with a laser transmitter (NEC, type NDL 5009) and receivers, a PIN diode followed by a transimpedance amplifier (ATT LG 1094).

Concerning the optical matrix, each hologram is 3 x 3 mm2 and is located at the nodes of a square array with a pitch of 6.35 mm; the glass plate is 10 x 10 cm2 (fig. 2).

From the purely optical point of view, one of the main difficulties to be overcome is related to transmission (precision on angular position of the emitter and divergence of the light beams). As the beams are pratically unguided in the glass plate, an initial angular offset can be very serious when a certain distance is reached. A precision of ± 10 mrd (at 1/e2) was actually obtained, which is fully compatible with the application. The beam divergence problem is also critical : a divergence of 10 mrd is an acceptable performance for integrated sources. This spreading of energy is followed by a reduction in energy reaching each hologram as the distance from the source increases, a phenomenon which is added with the reduction in energy related to the sampling of the previous receiving holograms (fig. 3).

As described, this technology has been demonstrated fully compatible with a broadcast diffusion of signals 9 * (1-->8), a figure that will be extended.

Each of the nine links was tested to the limits of the electro-optical interfaces used : at 1.4 Gbits/s with the NRZ code (or a bandwidth of 700 MHz). A BER of 5 * 10-11 was measured at 1.2 Gbits/s. The cross-talk between adjacent lines was measured and found to be less than 45 dB (optical). The dynamic between the 1st and 8th receiver on the same line was measured and found to be 23 dB, which is incompatible with receivers uniformity. This problem can be solved with the holographic matrix: Raising the number of etching operations and increasing the depth of the etching for the reception holograms as the distance to the emitter increases, results in a solution which quickly becomes very heavy and very costly. Fig. 3 suggests reducing the depth of etching for each hologram identically. The power received at the receivers is more uniform, but there is an additional constraint concerning receivers sensitivity and also the etching process which must be fully mastered.

4 - Conclusion

The use of holographic matrix offers an interesting alternative for solving broadcast problems of communication between circuit boards in the same rack. It allows to use architectures more matched to the large processors' quantities interconnected in massively parallel computers; the main advantage is the compatibility of a rise in frequency with a division of the signals. Holographic matrix backplane appears to be suitable for coarse-grain type computer's communication networks. However, for synchronous and massively parallel machines, this technology should be of great usefulness, especially for distribution of the clock and inter-processor synchronization signals.

The process used to produce holographic matrix employs means used by the semiconductor industry, but is less hampered by problems related to yield.An actual DRET program, specifically dedicated to multilevel optical matrix with electrically switched optical windows, must lead to increase the interest on this technology, allowing to make dynamic reconfiguration of a backplane and then to reach two major objectives: On the one hand, it offers a solution to failure recovery (redundant links allowing access to redundant processors elements). On the other hand, it allows functionnal reconfiguration on the network: the processors' arrangement can be modified, according to the application needs.

5 - References

(1) C. Sebillotte - "Holographic optical backplane for boards interconnections"
SPIE vol. 1389 "International conference on Advances in Interconnection and Packaging"

(2) C. Vergnolle, C. Sebillotte, B. Houssay, S. Paineau - "Optical Interconnection in airborne systems: Limitations due to backplane interconnections and prospects based on optical solutions" - L'Onde Electrique - vol.70 - n°6 - nov. 90

(3) C. Sebillotte - T. Lemoine - "Computer Generated Holograms directly etched in glass and their use as Boards Interconnects means" - IEE 3rd conference on "Holographic Systems, Components, Applications" - sept. 91

Massively Parallel Processing System with an Architecture for Optoelectronic Computing

Masatoshi Ishikawa, Akira Morita, and Nobuo Takayanagi

Faculty of Engineering, University of Tokyo
Bunkyo-ku, Tokyo 113, Japan
e-mail: ishikawa@k2.t.u-tokyo.ac.jp

I. Introduction

Recently there are growing interest in integration of optical devices such as OEIC's, surface emitting lasers, photo detectors, and so on. It is expecting that two-dimensional pattern information such as visual images can be treated in parallel. Parallel architecture using optical input/output devices and electronic processing elements (PE's) realize high speed parallel processing system with fully parallel architecture and parallel input/output. A conceptual diagram based on the architecture is shown in Fig.1. A hierarchical processing system by using optical interconnection as shown in Fig.1 and feedback type system in which output patterns can be fedback to the input of the system can be easily implemented.

Researches on integration of detectors and processing circuits have been done mainly from the viewpoint of integrated devices [1],[2]. C.Mead et al. [1] developed a silicon retina which realizes function of early vision integrating photo transistors, a resister network, and active circuits. MIT has vision chip project focused on analog integrated circuits with photo detectors [2].

However main circuits of these vision chip are fixed and has special purposes. A.Utsugi and M.Ishikawa [3] proposed a learning method of resister networks for adaptation of coordination system between a vision system and the real world.

On the other hand, considering that visual information is based on optical pattern, optical computing and learning capabilities of neurocomputing give effective tools for visual perception. M.Ishikawa et al. [4],[5] developed optical associative memory with learning capabilities and combined with conventional optical processing such as optical Fourier transform [6]. The system can carry out direct processing for the two dimensional visual information using optical parallel computing.

In order to realize flexible and high performance of parallel processing by using optoelectronic devices,

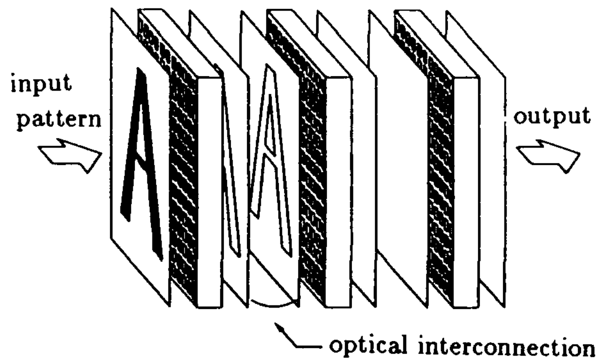


Fig. 1. Hierarchical optoelectronic processing system

combination of detectors and processing elements (PE's) is needed. In such parallel processing system to be integrated with photo detectors, a microprocessor in general can not be used as a PE of the system, because the number of gates in the microprocessor are too large that many PE's cannot be integrated into optoelectronic devices. Therefore parallel architectures using small scale PE's with generality of processing should be developed.

In this paper, an architecture for optoelectronic computing using massively parallel processing directly connected with photo detectors is proposed and an experimental system using originally designed and developed parallel processing LSI's with small scale PE's is shown. The experimental system as a scale up model of an optoelectronic parallel processing chip has a massively parallel processing architecture with 64x64=4096 PE's and the same number of photo detectors. Lastly, processing performance of the system is evaluated by carrying out some applications on the system.

II. System Architecture

A. Design concept

In this paper, the total system is regarded as a hierarchical parallel processing system as shown in

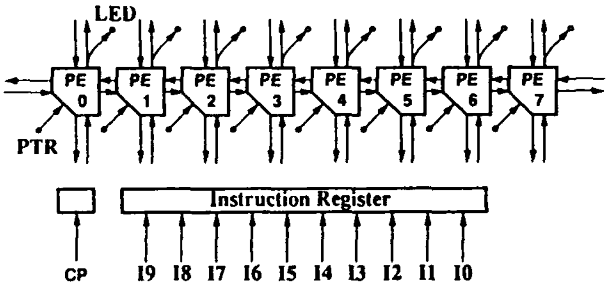


Fig. 2. Block diagram of SPE-8

Fig.1. The integrated vision system carries out detection and early stage processing of input data as a processing module of the lowest layer. Each layer is optically interconnected each other.

From the viewpoint of integrated optoelectronic parallel computing, an essential problem is how to implement processing modules with high generality into sensors. We propose an architecture in which pattern information is directly transferred to parallel PE's by using one-to-one interconnections. Since the architecture does not require serial transmission line of visual information as in the case of VIDEO signal, there does not exist an I/O bottleneck, therefore high speed processing is realized. However all circuits of processing elements including I/O interface should be integrated in a sensor device. In other words, the architecture aims at integration of massively parallel processing in practice.

In order to implement such integrated optoelectronic parallel processing system, the scale of PE should be compact. Although the compact PE has limitation of processing performance, parallel processing capabilities can compensate the limitation for some applications which have a computational structure matched with optical parallel processing. In this design concept, keys of the design are how generality of processing is kept by using compact PE's and how integration and high speed processing is performed.

B. Processing element (PE)

Although our final goal is to realize a general purpose optoelectronic device integrated with photo detectors, photo emitters and parallel PE's, as the first step, a compact PE is originally designed and eight PE's are implemented into an LSI. The LSI is named SPE-8 (Sensory Processing Elements - 8).

Block diagram of SPE-8 is shown in Fig. 2. The SPE-8 has eight PE's and each PE has input from a photo transistor (PTR) and output to a light emitted diode (LED), so perfectly parallel I/O and processing are implemented. All PE's are controlled by 10bit microinstruction through a single instruction stream.

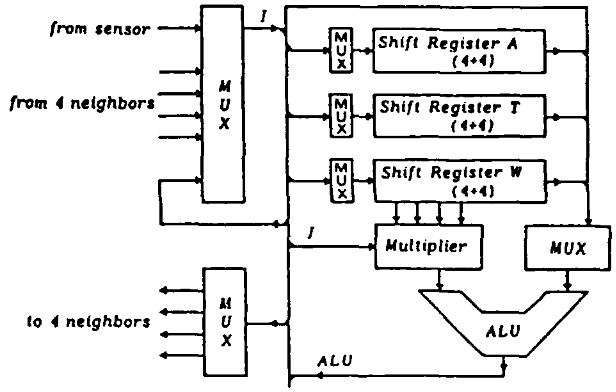


Fig. 3. Block diagram of processing element

Control mechanism of the system is effective for homogeneous data parallel processing such as pattern processing and scale of processing elements can be so compact. Each PE is interconnected with the 4 neighborhoods. The SPE-8 is designed to be cascadable with arbitrary number of SPE-8's.

Each PE has three 8bit registers (A:Accumulator, T:Template, W:Weight), one arithmetic logical unit (ALU, 1bit), one 4bit multiplier as shown in Fig. 3. Processing architecture of the PE is based on bit serial processing which is slow in comparison with bit parallel processing, but dose not require so many number of gates, then it has major advantages for integration or variable bit length processing.

Functions of ALU include AND, OR, Exclusive OR, addition, subtraction, multiply(4bits×4bits) and combinations of these basic functions such as weighted sum for calculation of correlation.

The instruction set of the SPE-8 is rearranged from 32bit direct horizontal type microcode to four kinds of 10bit microinstruction by classifying the instructions. The instructions except instructions to carry out active processing are held in instruction registers of the SPE-8 until the next change of the instructions, because these types of instructions are almost always fixed for the period of the active processing. By this method, not increasing the number of instructions in practice, equivalent control with horizontal microprogramming is realized by small gates of the PE and small bit length of the instruction.

In the result, the PE is implemented by using 337 gates, then eight PE's (8 × 337 = 2696 gates) and a common part (274 gates) implemented in one chip of CMOS gate array (max. 3312 gates). Processing cycle time is estimated typ. 44ns, max. 87ns by a simulation of critical pass (load capacitance 30pF). All functions of the LSI have been tested by logic

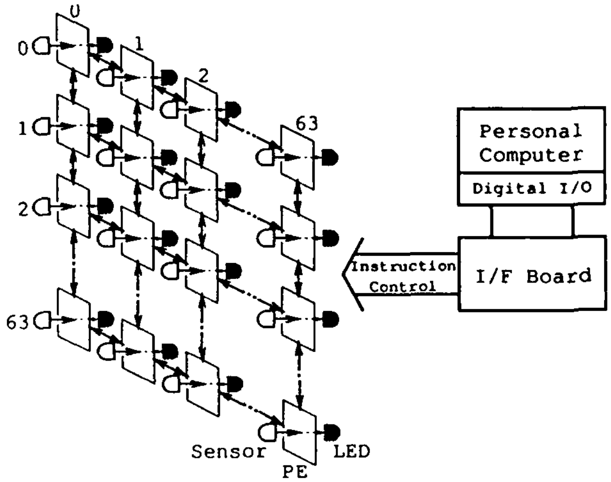


Fig. 4. Block diagram of SPE-4k

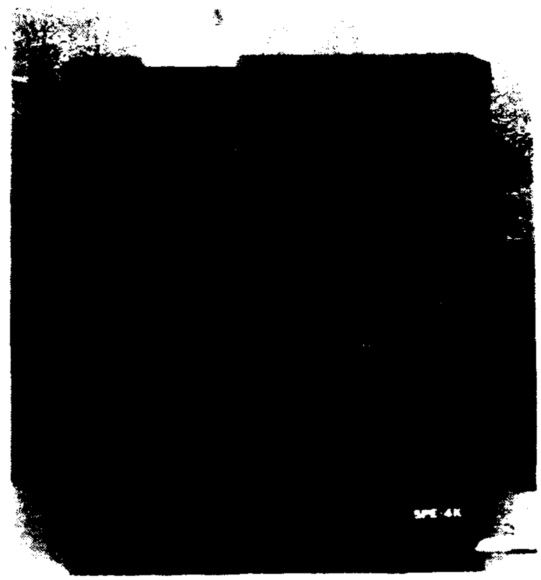


Fig. 5. SPE-4k system

simulation, by LSI testers and by in-circuit test.

C. SPE-4k

An experimental optoelectronic processing system which has matrix positioned $64 \times 64 = 4096$ (4k) PE's is constructed by using 512 SPE-8's. The system is named SPE-4k and is seemed to be a scale up model of integrated optoelectronic processing elements. The structure of SPE-4k is shown in Fig. 4 and an overview of SPE-4k is shown in Fig. 5. PTR's are positioned with the space of 12mm. LSI's are arranged between the front LED array and the back PTR array.

Two kinds of control method are designed. One is a method by using microprogramming control in which the control computer may send macro instructions to be decoded by a microprogram sequencer. The other is a method by using I/O interface of the control computer.

In this paper, the SPE-4k system uses I/O processor type of control in order that verification of processing behavior and development of software may be carried out on the computer (Fig. 4). However the processing speed is limited by the speed of parallel I/O of the computer. The cycle time of the system is 10μs. As the SPE-8 itself can work in 100ns cycle time, so the speed of the present system is 1/100 of maximum speed of the SPE-4k. Considering the 8bit integer addition as a basic operation of processing for the evaluation of the speed of SPE-4k, total 32MOPS (Mega Operations Par Second) by the present system and 3.2GOPS (Giga Operations Par Second) at the maximum speed of the system are obtained. The processing speed is very high speed in comparison with general image processors or parallel processing systems.

III. Applications

In order to evaluate processing performance of the SPE-4k, the following applications have been carried out on the system. The processing time of the following applications is shown in Table I. The column of maximum speed shows performance at maximum speed of the SPE-4k (cycle time: 100ns).

As shown in Table I, high speed processing in order of micro seconds is obtained. The processing speed is not limited by VIDEO rate (16.7ms), so it is about 10^5 times faster than that of conventional image processors which use the VIDEO signal

A. Edge detection

Two and four neighborhood edge detection to the 1bit PTR input pattern p_{ij} are implemented by logical operation. Operation of two neighborhood type edge detection is

$$q_{ij} = (p_{ij} \div p_{i-1,j}) \cup (p_{ij} \div p_{i,j-1}), \qquad (1)$$

and four neighborhood type edge detection is

$$\begin{aligned} q_{ij} = {} & (p_{ij} \div p_{i-1,j}) \cup (p_{ij} \div p_{i,j-1}) \\ & \cup (p_{ij} \div p_{i+1,j}) \cup (p_{ij} \div p_{i,j+1}) \end{aligned} \qquad (2)$$

where $\div$ and $\cup$ are EXOR and OR respectively, and q_{ij} is output pattern.

B. Skeletonization

Simple skeletonization using 1bit input data and

TABLE I
PROCESSING TIME OF APPLICATIONS

processing	no. of steps	processing time present system	maximum speed
edge detection			
(2 neighbors)	23	230μs	2.3μs
(4 neighbors)	33	330μs	3.3μs
skeletonization			
(2 neighbors)	149	14.9ms	149μs
(4 neighbors)	360	36.0ms	360μs
detection of moving objects	7	70μs	0.7μs
trace	4	40μs	0.4μs
Poisson's equation	125	250ms	2.5ms

four or eight neighborhood data is implemented. Preparing neighborhood patterns of edge, pattern matching with input data and judgement of edge are carried out. Ten times iteration of this operation is made on the SPE-4k. Two and four matching pattern are prepared for four neighborhood and eight neighborhood, respectively. Rotated patterns are also matched with the templates.

C. Detection of moving objects

Edge detection of moving objects using time derivative of 1bit input data. The operation is

$$q_{ij}(t) = p_{ij}(t) \oplus p_{ij}(t - \triangle t), \tag{3}$$

where $p_{ij}(t)$ and $q_{ij}(t)$ are PTR input and output at the time t, respectively. Sampling time of detection $\triangle t$, which is determined by velocity of the objects, is available to 0.7μs at the maximum speed of the SPE-4k.

D. Trace

Trace of moving objects using logical OR operation between memory data and temporal input data is implemented. The operation is

$$q_{ij}(t) = p_{ij}(t) \cup q_{ij}(t - \triangle t). \tag{4}$$

This operation is OR between input $p_{ij}(t)$ and internal memory just one step before, where $\triangle t$ is 0.4μs at maximum speed of the LSI.

E. Solution of Poisson's equation

As an example of solution of simple partial differential equation, two dimensional Poisson's equation $\nabla^2 u = -f$ is solved in 8bit fixed point operation for 1bit PTR input patterns. Since the solution is obtained as a converged pattern of a diffusion equation related to the Poisson's equation:

$$\nabla^2 u = c\frac{\partial u}{\partial t} - f, \tag{5}$$

iterations for the conversion is needed. In addition, Eq.(5) is changed to a difference equation:

$$u_{ij}^{n+1} = \frac{1}{8}(u_{i-1,j}^n + u_{i+1,j}^n + u_{i,j-1}^n + u_{i,j+1}^n) + \frac{1}{2}u_{ij}^n + \frac{1}{8}f_{ij}. \tag{6}$$

The solution is obtained after 200 iterations.

IV. CONCLUSION

Massively parallel processing system SPE-4k using an architecture for optoelectronic computing is shown and the speed and processing capabilities are shown by carrying out some basic applications. The system has $64 \times 64 = 4096$ PE's and the same number of PTR's.

In the SPE-4k, compact processing elements realize general early vision processing by using bit serial architecture. Since the architecture requires small number of gates in one PE, the system will be able to be integrated into one chip optoelectronic devices in future.

REFERENCES

[1] C.Mead, *Analog VLSI and Neural Systems*, Addison-Wesley, 1989.

[2] J.L.Wyatt, et al., "The MIT vision chip project: analog VLSI system for fast image acquisition and early vision processing," *Proc. 1991 IEEE Int. Conf. on Robotics and Automation*, pp.1330-1335, 1991.

[3] A.Utsugi and M.Ishikawa, "Construction of inner space representation of latticed network circuits by learning," *Neural Networks*, Vol.4, No.1, pp.81-89, 1991.

[4] M.Ishikawa, et al., "Optical associatron: a simple model for optical associative memory," *Appl. Opt.*, Vol.28, No.2, pp.291-301, 1989.

[5] M.Ishikawa, et al., "Experimental studies on learning capabilities of optical associative memory," *Appl. Opt.*, Vol.29, No.2, pp.289-295, 1990.

[6] M.Ishikawa, et al., "Optical Associative Memory Combining with Optical Preprocessing," *Technical Digest of OPTICAL COMPUTING '90*, pp.160-161, 1990.

A Multiprocessor System Using Interboard Free-Space Optical Interconnects: COSINE-III

Toshikazu Sakano, Kazuhiro Noguchi, and Takao Matsumoto
NTT Transmission Systems Laboratories
1-2356 Take, Yokosuka-Shi, Kanagawa-Ken, 238-03 Japan.
TEL:+81-468-59-3194, FAX:+81-468-59-3396.

1. Introduction:

Optical interconnects are expected to play an important role in multiprocessor systems of the future[1]. This is because they have many advantages over conventional electrical interconnects, such as high speed, wide bandwidth, and immunity from electromagnetic interference. In addition to these advantages, free-space optical interconnects offer the unique feature that signal transmission can be realized without any 'physical' contact[2]. This feature makes it easy to construct three-dimensional interconnection networks in all the interconnection layers of a system; from chip-to-chip to frame-to-frame.

In most multiprocessor systems, interboard(board-to-board) interconnects are usually realized on the back plane and leads to several critical drawbacks such as wiring congestion, increased signal delay and variation (clock skew), as the number of processors and clock speed are increased. To overcome these drawbacks, some multiprocessor systems that replace the back plane interconnects with free-space optical ones have been reported[3],[4]. However, no system yet reported fully utilizes the feature that interboard optical connections can be set up any where on the board to achieve the shortest path length, except one paper describing optical equipment for such use[5].

This paper describes a multiprocessor system implemented using the interboard free-space optical interconnect scheme named COSINE-III(Computer system employing Optical Spatial INterconnections for Experimentation-III). The system interconnects 64 processing units in a three-dimensional mesh network through 48 bi-directional free-space optical interconnects which are distributed on both sides of four processor boards and several electrical links. The configurations of the system and the free-space optical interconnects are described. Some preliminary test results are reported.

2. System Configuration:

COSINE-III is a multiprocessor system that has a three-dimensional mesh processor network. The network was chosen for the reasons that the configuration conceptually and structurally complements the strengths of free-space optical interconnects, and that the network is simple and more suitable for various scientific applications than other static processor networks. The configuration of COSINE-III is schematically shown in Fig.1. The system is comprised of four stacked processor boards, each of which has 16 processing units that are arranged in a two-dimensional 4×4 grid. Each processing unit has six bi-directional communication links to connect it to its adjacent processing units. Four of them are electrical interconnects which connect the four adjacent processing units on the same board. The two remaining links are free-space optical interconnects to the two immediately adjacent processing units on neighbouring processor boards. As shown in Fig.1, the free-space optical interconnects make it possible to provide direct communication between the processing units on adjacent boards without any 'physical' wires. The three-dimensional nature offers such advantages as (1) avoidance of wiring congestion and crosstalk, (2) small signal delay and variation, and (3) increased flexibility for system installation.

3. Free-space optical interconnects:

In COSINE-III, the free-space optical interconnects link the processor boards as shown in Fig.1. This configuration, however, does permit many environmental factors to impact the signal transmission characteristics. They are (1) fluctuations in detected optical power caused by board displacement and/or vibration, (2) leakage of environment light, (3) variation in coupling efficiency between the optical sources and the detectors caused by device fabrication inaccuracy and (4) DC drift of electrical and optical devices caused by heat. The interconnect signal transmission scheme must, therefore, not only be reliable enough to overcome the environmental factors, but also simple so as not to waste the board area. To meet these requirements, COSINE-III employs the differential signal transmission method[6].

Fig.2 shows the configuration of the differential signal transmission method for the free-space optical interconnects in COSINE-III. The driver on board S, modulates the optical sources LED-1 and -2 with the input TTL signal and its inverted equivalent, respectively. At the receiver on board R, the two optical beams are detected by detectors PD-1 and -2, respectively. Their differential signal is amplified and reshaped by the comparator. The reshaped signal is converted to a TTL signal by the level converter and output from the receiver.

With the configuration described in Fig.2, the environmental factors are common to the two optical channels and are canceled in the process of making the differential signal. That is, the optimum threshold level in the decision process is always the same regardless of the envirnomental factors as long as they are common to the two optical channels. Furthermore, since the method is equivalent to bipolar signal transmission, the optical power at each detector necessary to achieve a specified bit error rate(BER) is about 3dB less than is needed in the conventional unipolar optical signal transmission method.

4. Fabrication and preliminary test:

Prior to constructing a prototype system, a free-space optical interconnect was fabricated and its characteristics were evaluated. The specifications of the fabricated interconnect are summarized in Table 1. An LED (λ=0.8 mm) and a collimating lens were combined in a LED module. A PD module was formed by joining a focusing lens to a PIN photo diode. Two LED modules were mounted on a board as closely as possible (8 mm pitch). Two PD modules were fixed to another board with the same spacing. These two boards were placed face-to-face with a separation of L. The peak optical power output from both LED modules was +2 dBm at a drive current of 20 mA. The average received optical power at the facing PD module was -1.5dBm when L = 5 cm. With this optical setup, BER characteristics at 20Mb/s were measured first. The average received optical power needed to achieve a BER of 10^{-9} was -26.7 dBm. The value was far lower than the typical received power of the setup, and was 2.3 dB lower than that needed by the conventional unipolar signal transmission method. Next, the relationships between BER, L and lateral board displacement d were measured. The results are shown in Fig. 3. With -2.5mm < d < 2.5mm, the BER was under 10^{-11} for L=1cm to 7cm. It is therefore concluded that the tolerance for lateral board displacement is ±2.5mm and it is possible to achieve such accuracies with existing fabrication technologies.

A prototype COSINE-III system was designed and fabricated based on the results of the free-space optical interconnect preliminary test. The specifications of the fabricated system are summarized in Table 2. Each processing unit was formed by cascading two Transputers(T800s), each of which has four bi-directional communication links and 4 MByte of local memory. One processor board accommodates 16 processing units which are arranged in a two-dimensional grid and 32 bi-directional free-space optical interconnects (16 interconnects on each side of the board). Conventional epoxy-glass boards (438mm × 450mm × 1.6mm) were used. The four processor boards were accommodated in a machine frame. Fig. 4 is a view of the fabricated system with a side panel removed. The outside dimensions of the system are 500mm wide, 760mm high and 460mm deep. The spacing between boards in the frame is

25mm. Fig. 5 is a closeup of one processor board. The transputers and associated memories, which are the square chips and the densely stacked chips, respectively, are arranged two-dimensionally on both sides of the board. LED modules, PD modules are attached to the board with stiffeners to avoid misalignment due to board warp. In the photograph, the modules are the rows of small dots between the processors.

The fabricated system was tested to confirm signal transmission through the free-space optical interconnects. Fig. 6 shows the histogram of signal delay time t_d for all 96 optical interconnects. The average delay time was 18 ns. Most of the delay was attributable to the electrical circuits like the LED driver, the comparator, and TTL buffers, and could be reduced by refining these circuits. Some interconnects had a relatively large delay time. These delays were caused by offset deviations in the comparators and could reduced by carefully matching the comparators. Signal transmission was successfully established even after repeated extraction and insertion of the boards. Next, a program that performed continuous random data communication among 8 processing units through 4 bi-directional free-space optical interconnects and 4 electrical ones was loaded and run on the system. It was confirmed that the program ran without error for 100 hours. It was therefore concluded that the fabricated system is stable enough to use as a multiprocessor system.

5. Summary:

A multiprocessor system using interboard free-space optical interconnects, COSINE-III, was constructed. The system forms 64 processing units into a three-dimensional mesh network with the help of 48 bi-directional free-space optical interconnects distributed on both sides of 4 processor boards. We confirmed that the fabricated system worked well as a multiprocessor system.

References:

[1] Special issue on supercomputing, *IEEE SPECTRUM*, vol. 29, No. 9, pp. 69-75 (1992)
[2] J. W. Goodman, et al., *Proc. IEEE*, vol. 72, pp. 850-866 (1984)
[3] T. Matsomoto et al., *Proc. ICCD'90*, Cambridge, pp. 426-429 (1990)
[4] J. W. Parker, in *Tech. Digest Optical Computing '91*, Salt Lake City, pp. 304-307 (1991)
[5] Dean Z. Tsang, in *Tech. Digest Optical Computing '91*, Salt Lake City, pp. 132-135 (1991)
[6] P. J. Ayliffe, et al., *SPIE*, vol. 1281, pp. 2-15 (1990)

Table1 Specifications of a free-space optical interconnect

Optical source	LED (wavelength:0.8μm)
Collimating lens	Planer convex (D=5mm, f=3.1mm)
Output power	+3dBm @ drive current; 20mA (Typical)
Detector	Si-PIN photo diode Detection region:0.9mm square
Focusing lens	Planer convex (D=5mm, f=3.1mm)
Received power	-1.5dBm @ board spacing; 5cm (Typical)
Module spacing	8 mm (10mm in COSINE-III)
Transmission speed	20Mb/s

D;diameter, f;focal length

Table2 Specifications of COSINE-III

Processors	Transputer(T800)
Local memory per processor	4MByte
Number of PUs per board	4×4
Number of boards	4
Network configuration	3-dimensional mesh
Network scale	4×4×4
Board spacing	25mm
Board dimension	438mm×450mm×1.6mm
Frame dimension	500mm×760mm×460mm
Bi-directional optical links	48 (16×3)

PU;Processing Unit comprised of two cascaded processors.

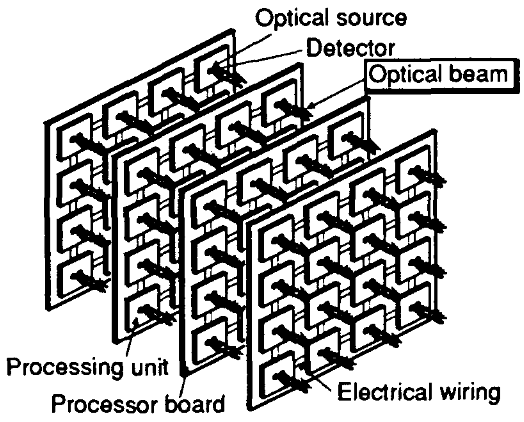


Fig.1 System configuration of COSINE-III

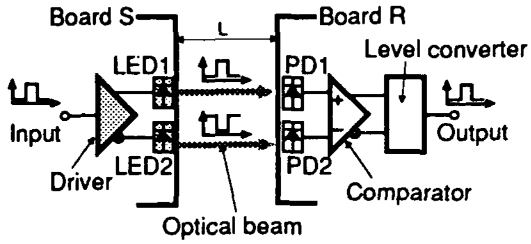


Fig.2 Differential signal transmission method for the free-space optical interconnects.

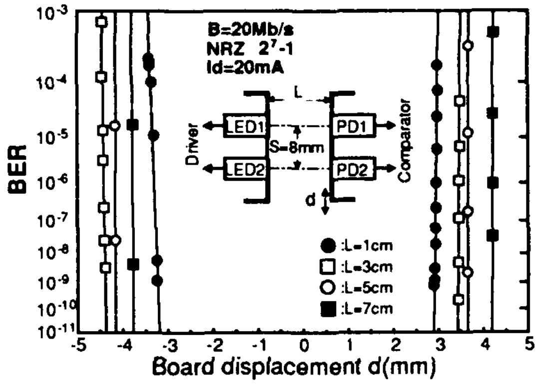


Fig.3 Relationship between the board displacement and bit-error-rate

Fig.4 Photograph of the fabricated system. A side panel is removed.

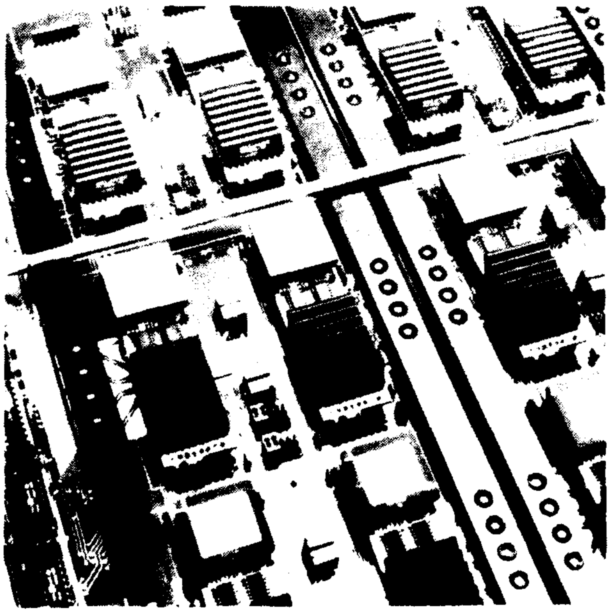


Fig.5 Photograph of a processing unit and related free-space optical interconnects on a board. The small spots arranged in a row on the black stiffeners are the LED/PD modules.

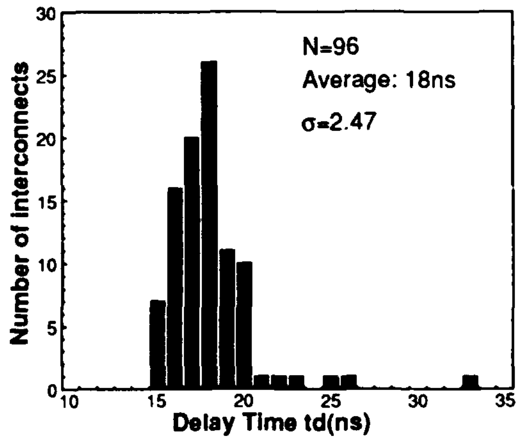


Fig.6 Signal delay time of free-space optical interconnects in COSINE-III

An Optically Connected Parallel Machine

D. M. Monro*, J. A. Dallas*, J. A. Nicholls*, M. D. Cripps† and W. A. Crossland§

*School of Electronic and Electrical Engineering, University of Bath, Bath, BA2 7AY, England
Telephone: +44-225-826833 Fax: +44-225-826073

†Department of Electrical, Electronic and Information Engineering, City University,
Northampton Square, London EC1V 0HB Telephone: +44-71-477-8134 Fax: +44-71-477-8568

§Engineering Department, Cambridge University, Trumpington Street, Cambridge, CB2 1PZ,
Telephone: +44-223-330264 Fax: +44-223-332616

1. Introduction

Supercomputers are entering a phase in which the quest for faster communications in multiprocessor systems will be a dominant issue[1]. Over many years the achieved performance of advanced computers has become problem dependent, as a result of architectures which attempt to reconcile high computing performance with low communication bandwidth. SIMD machines have very large numbers of simple processors operating in synchrony on a shared memory. These achieve peak performances close to their ratings on problems which are genuinely data parallel, but they are ill-suited to problems which are naturally expressed in terms of code parallelism. MIMD machines have used smaller numbers of processors running independent programs which pass messages, and although difficult to program, these have succeeded in problems which require asynchronous processors to exchange relatively long messages, relatively infrequently.

Recent developments indicate a blurring of the distinction between SIMD and MIMD architectures. Whether or not memory is actually shared is becoming less obvious to the user, as computers with distributed physical memory present a shared memory model to the user, supported by software. A number of recent and projected computers have moderately large arrays of RISC processors, such as might formerly have been used in MIMD machines. Each executes the same program as in a SIMD computer, but asynchronously and with data not actually shared. To succeed, these machines will require high communication bandwidths which scale with the number of processors. By implication, if processors are logically or physically grouped into clusters, the bandwidth of intergroup communication will have to be very high indeed.

The ideal communications network would exchange items of information among processors as fast as they compute with them. This requires a full crossbar capability between the processors in such a system, and large bandwidth communications. Until recently this might have seemed an unrealistic demand, but optical technology offers genuine possibilities. It is well known that optical space switches have the crossbar capability, and that optical communications technology can approach Terrabit transmission rates per channel. Full single stage crossbars can be built that are strictly non-blocking, and capable of full broadcast and multicast. In addition, optical interconnects provide complete electrical isolation, and they neither radiate, nor are affected by, electromagnetic radiation.

Processors with 64 bit wide processing running at 100MHz are in prospect, and these ought ideally to be able to communicate in a shared memory system at this bandwidth, i.e. 6.4 Gigabits/Sec per processor to achieve transparency in memory sharing. This demand is mitigated to some degree by the use of a cache, and we will describe a feasible approach to such a system. The combination of a 64x64 matrix-matrix LCD Spatial Light Modulator (SLM) for switching with the use of fibrechannel technology for data transmission comes closer to this ideal than any feasible electronic network, with the only current deficiency being in its switching speed.

Large all-electronic single stage full crossbar switches are not practical because of the cost associated with large numbers of crosspoint elements and the connectivity problems of 2-dimensional electronics. As the required channel bandwidth and the number of processors increases, the complexity of electronic switches becomes formidable.

Serial electronic communication and switching, whist practical and convenient, is restricted in bandwidth by comparison with our optical approach. To achieve the required bandwidth, at least a byte-wide path must be switched, and the pin limitations of packaging technology would limit the size of crossbar which could be contemplated. With interconnections such as HIPPI, this pin limitation quickly becomes the dominant design factor. Freespace optical switching avoids this restriction.

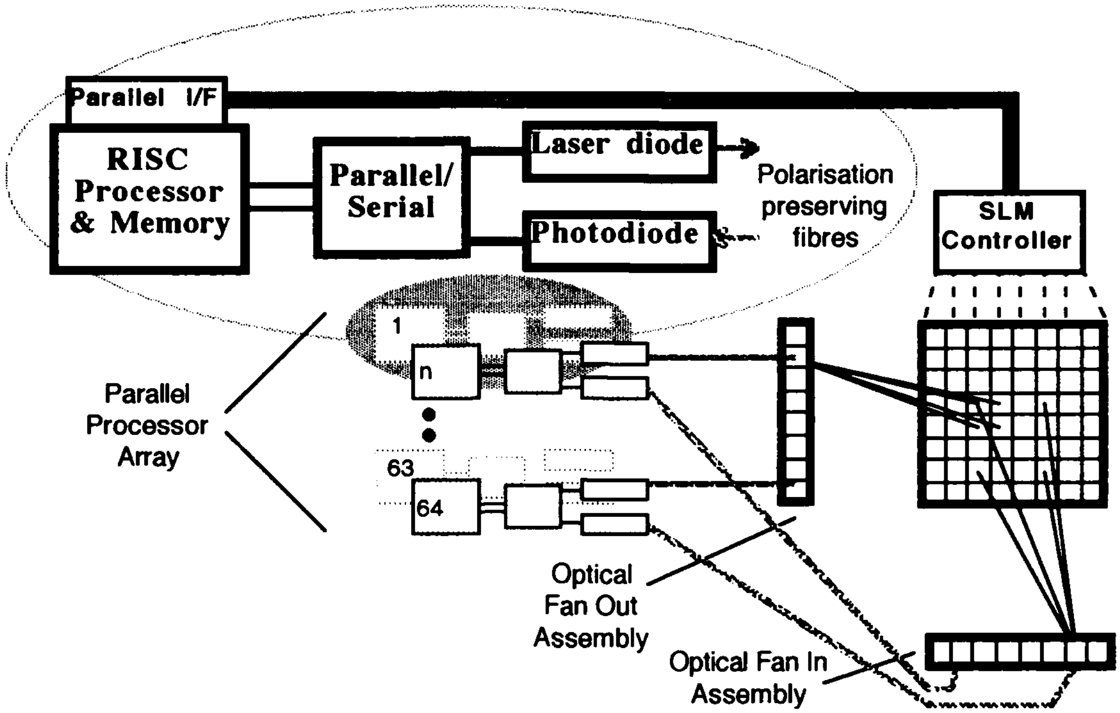


Figure 1. The architecture of the Optically Connected Parallel Machine (OCPM) showing an exploded view of a typical processor and its interaction with the switch.

2. The OCPM

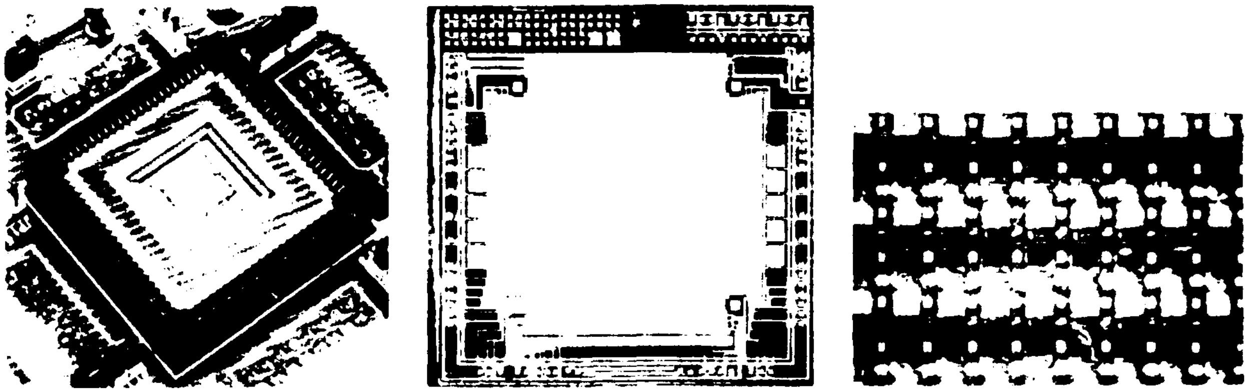
The OCPM Project (Optically Connected Parallel Machines) is a collaboration between industrial and academic partners in the United Kingdom: British Aerospace (Sowerby Research Centre), Meiko, BNR Europe, Heriot-Watt University, The University of Bath and Thorn EMI CRL. The project is coordinated by British Aerospace and is funded in part by the DTI and SERC.

The architecture of the OCPM is illustrated by Figure 1. The machine will use state of the art Sparc RISC processors. Each processor has two 10-bit wide 70 M Hz communication links. One of these will be used to communicate with the switch controller, and the other will carry the data. It is proposed to carry out parallel/serial and serial/parallel conversion and diode modulation using standard fibrechannel chip sets. To maintain polarization onto the SLM, laser diodes custom pigtailled to polarization preserving fibres will be used. In addition, custom pigtailing of multimode fibres to the photodiodes will be employed to reduce losses.

Optical crossbar switches based on the vector-matrix multiplier principle [2] have been described by a number of authors [3]. The crossbar switch in the OCPM is based on the matrix-matrix principle [4,5], which offers enhanced scalability, performance and compactness [8]. It is coupled by polarization preserving fibres to a compact free space optical system described elsewhere at this meeting[9].

The SLM switch is an array of shutters or mirrors, based on a ferroelectric liquid crystal layer directly over a silicon VLSI die [8, 9, 10]. Circuitry in the silicon backplane controls and operates the ferroelectric liquid crystal shutters. SLMs of this type have been built with arrays of 176x176 switchable mirrors on a 10mm square silicon die [7]. Figure 3 shows such an array in which each mirror might be considered as an optical crosspoint. The SLM to be constructed for the OCPM will be a 64x64 array. With presently available liquid crystals and utilising graded junction drive transistors on a 3.5 mm CMOS silicon backplane, switching times of 10 μsec or less might be achieved.

In the OCPM the initial implementation will use an external electronic network to control the switch, and much of the control circuitry will be external to the SLM chip. Later versions could integrate this functionality into the silicon SLM backplane [7], with a consequent gain in speed of the electronic control. Fibrechannel protocols will be used as the physical and signal layers for communication, taking advantage of the emergence of standards and the economies inherent in them. Arbitration with the switch will be implemented in software,



a) A fully assembled 176x176 pixel array device

b) The silicon substrate for the above SLM. The backplane size in 10x10 mm

c) A photograph showing even rows of the pixels switched on

Figure 2. A silicon backplane ferroelectric liquid crystal SLM with 176 by 176 switchable pixel mirrors. The device is built over a 10 x 10 mm silicon die. The OCPM will use a similar 64 x 64 pixel device.

both in the switch controller and the processing nodes. A processor will initiate a channel request and pass it to the switch controller. Because of the minimum 10 μsec latency as a liquid crystal shutter settles, processors will be free to do other work until the switch controller acknowledges the availability of a channel for transmission. Controller arbitration will take about 0.5 μsec per request in addition. A single isolated request would therefore be met in 10.5 μsec, but reconfiguration of all 64 channels could take 42 μsec for the SLM to become available on all channels.

3. Performance evaluation

We have simulated the operation of the OCPM as a communications network. We assume that processors demand random connections at normally distributed time intervals, and then send one packet. Channels are acquired on a first-come first served basis. The simulated loading of the network is increased by reducing the mean time between requests. The simulation applies the arbitration and switching latencies, and we measure the throughput per channel assuming that communications proceed at 640 MBits/sec.

To send a message, the simulation

1. Waits for the transmitting processor to finish sending any other messages (processor wait latency).
2. Waits a certain time for arbitration.
3. Opens the switch to its destination, with 10 μsec settling time.
4. Waits for a clear path to its destination (channel wait latency).
5. sends the message at 640 M Bits/sec.

Figure 3 shows the simulation results with 8K bit packets, where 20% of the messages passed were 8-way multicasts, i.e. fanned out to 8 receivers. At low loads, the latency is dominated by the channel wait latency. At higher loads (70% of maximum), throughput begins to saturate, because messages build up in the processor waiting to be sent, and the processor wait latency explodes, degrading performance dramatically.

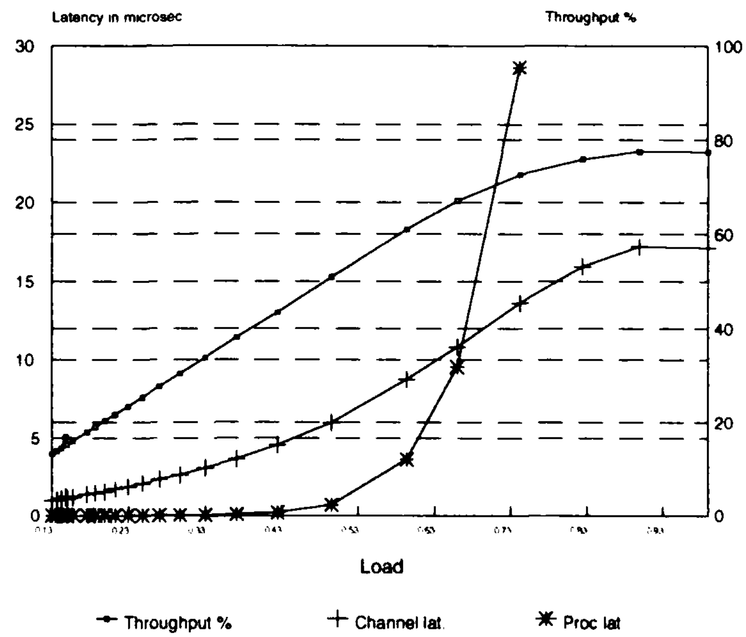


Figure 3. Behaviour of latencies in the OCPM showing breakdown of performance at 70% of theoretical maximum throughput.

Figure 3 suggests that an OCPM using multicasts, as in this simulation, is best run at average loads of lower than 60%, i.e. below 384 M Bits/sec per processor.

Figure 4 compares the total message latency as a function of throughput, for simulations with different packet sizes, different multicast probabilities, and different arbitration times. It shows the performance with 8K bit packets to be better that 16 K bit packets. This is because the time taken to transmit an 8K packet is 12.8 microseconds, which approximates the switching time. We also see the useful advantage that multicast brings to the throughput. Finally one simulation uses a 5 microsecond arbitration, delay which is intended to indicate a lower bound on performance for 8K packets.

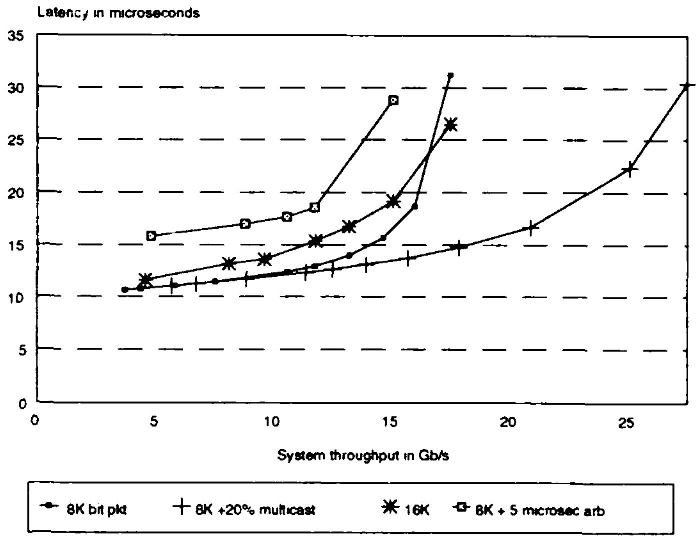


Figure 4. Some comparisons of OCPM performance.

4. Conclusions

Electronic switching of 150 M bits/sec channels in 85 nsec was demonstrated by one of us in 1985 [10]. Advances in design and technology now permit sub 50 nsec switching of more than 500 M bit/sec channels in up to 16 by 16 crossbars. Such electronic technology is approaching its limits. The 64 by 64 optical SLM described here gives this order of performance at the start of its development cycle.

The current restrictions on switching sppeds and inline switching of freespace optical mechanisms such as the one used in the OCPM are not fundamental. Improved layout and materials could improve the switching speed by at least an order of magnitude. In addition, solutions are possible which would permit inline addressing using an extra matrix element to decode the optical stream and electronically switch within the same SLM substrate [7].

In conclusion, therefore, the OCPM demonstrates a feasible approach to easing the communications bottlenecks experienced in multiprocessor systems. The OCPM consortium plans to construct a partially configured functional Optically Connected Parallel Machine by 1994 operating as described in this paper.

5. References

1. Bell, Gordon, 'Ultracomputers, a terraflop before its time', Comm ACM, Vol. 35, No. 8, pp. 27-47, August 1992.

2. Sawchuck, A A, Jenkins, B K, and Raghavendra, C S, Computer, pp. 50-60, June 1987.

3. Dias, A R, Kalman, R F, Goodman, J W, and Sawchuck A A, Opt. Eng., Vol. 22, No. 1, pp. 955-960, 1988.

4. De Groot, P J, and Noll, R J, Appl. Opt., Vol. 28, No. 8, pp. 1582-1587, 1989.

5. Kirk, A G, Crossland, W A and Hall, T J, Proc SPIE, 1574, 1991.

6. Collings, N J, Crossland, W A, Aycliffe, P J, Vass, D G and Underwood, I, Appl. Opt. Vol 28, No. 2, pp. 4720-4747, 1989.

7. Crossland, W A, Birch, M J, Davey, A B and Vass, D G, SPIE/IS&T Symposium on Electronic Imaging Science & Technology, paper 1665-36, 1992.

8. Crossland, W A, Birch, M J, Kirk, A J, Hall, T J, Vass, D G, OSA Topical Meeting on Smart Pixels, 1991.

9. White, H J, Brownjohn, N A, Stace, C, Proudley, G M, Walker, A C, Taghizadeh, M R, Robertson, B and Barrett, C P, and Birch, M J, 'Practical demonstration of a free space optical crossbar switch' , Topical Meeting on Photonics in Switching 1993, Optical Society of America, March 1993.

10. Field, A J, and Cripps, M D, 'Self-clocking networks', Proc. IEEE Int Conf Parallel Processing, pp. 384-387, Chicago, August 1985.

Cellular processing with diffractive optical elements

Andrew Kirk, Tomohira Tabata and Masatoshi Ishikawa
Faculty of Engineering, University of Tokyo
Bunkyo-ku, Tokyo 113, Japan
Tel: +81 3 3812 2111 ext. 6902

I. Introduction

As the number of nodes within a parallel processing system is increased it becomes more difficult to achieve a global interconnection between them. There are however many examples of massively parallel processing systems for which global interconnections are not required. Early vision processing operations such as edge detection and feature recognition can be performed by the use of a single interconnection pattern which is repeated in a shift-invariant manner over an entire array of cellular processing elements. Mead [1] developed a silicon retina in which neighbourhood interconnections were performed electronically. More recently Ishikawa has developed a high-speed programmable system for early vision processing (SPE-4K) [2]. This system consists of an array of 64 x 64 processing elements (PEs) each of which can accept optical input via a photodetector and provide optical output via a LED. Each PE contains an arithmetic and logical unit (ALU) which can perform 8 bit processing operations and which is electronically interconnected to the 4 neighbouring PEs. The instruction cycle time is 100 ns and operations such as edge detection can be performed in 3.3 μs . The architecture of this system has been designed in such a way that in future it may be fabricated as single integrated device. This system may thus be thought of as a scaled-up model of an integrated high-speed optoelectronic parallel processing system.

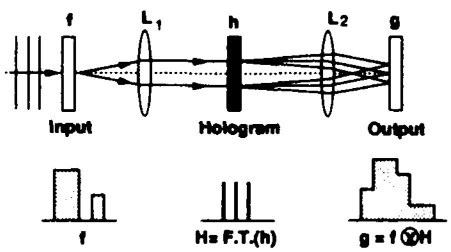


Figure 1: Convolution operation with a Fourier plane hologram.

Although nearest-neighbour interconnections are relatively simple to implement electronically the interconnection density at each PE increases rapidly with the interconnection neighbourhood size. This has motivated us to investigate the application of Fourier plane holographic interconnections to provide the interconnection for the SPE-4K system. The use of Fourier plane holograms in this way was first introduced by Jenkins et al [3] and has since been extended other authors [4,5]. One major disadvantage in using CGHs in this way is that they result in a fixed interconnection and so the interconnection pattern cannot be modified. In this paper we introduce a new system which allows reconfigurable shift-invariant interconnections between large arrays of PEs of the type described above. The scalability of this system will be assessed and preliminary experimental results will be presented.

II. System design

The basic principles of this scheme are illustrated in Fig.1. The optical system performs a convolution of the input pattern f with a convolution kernel determined by the impulse response H of the hologram. This kernel is given by the optical Fourier transform of the hologram transmission function h, i.e. $H = \mathcal{F}(h)$ where $\mathcal{F}$ represents the Fourier transform. Thus the output g is given by $g = f \otimes H$. The convolution operation in Fig.1 is thus equivalent to the shift-invariant mapping of the input plane to the output plane with nearest-neighbour interconnections. The hologram defines this mapping simultaneously over the entire processing plane. If the input pattern f consists of an array of mutually incoherent sources (such as a LED array) then the output intensity at the exit plane is given by the incoherent convolution of the input intensity with the modulus of the hologram impulse response (i.e. $|g|^2 = |f|^2 \otimes |H|^2$). This requires that only the intensity response of the CGH ($|H|^2$) is determined and so greatly simplifies CGH design and fabrication.

The optoelectronic parallel processing system is shown in Fig.2. Each processing element (PE) consists of a photodetector connected through an ALU to a LED as described earlier [2]. The interconnection system essentially a folded version of Fig.1 in

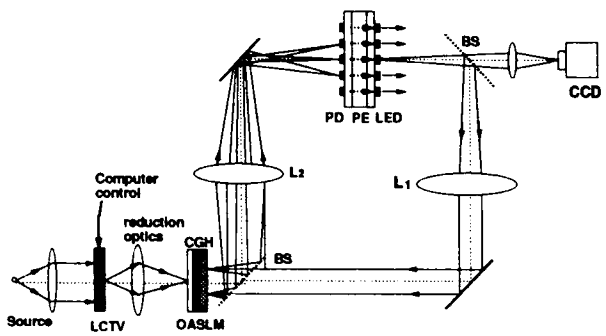


Figure 2: Shift-invariant optical parallel processing system (PD - photodetector, PE - processing element, BS - beam splitter).

which the output pattern is imaged onto the detector plane with holographic shift-invariant interconnection. In contrast to the system shown in Fig. 1 the hologram employed here is reconfigurable A binary CGH function is calculated and is displayed on the liquid crystal television (LCTV) as a periodically replicated pattern. This is then imaged onto the ferroelectric liquid crystal (FELC) optical addressed spatial light modu'ator (OASLM) so that the CGH function is copied into the phase transmittance function of the OASLM. The OASLM thus acts as a reconfigurable phase CGH. Reconfigurable CGHs have been implemented previously by the use of an electrically addressed SLM (EASLM) [6]. At present OASLMs have a significantly better resolution ($\simeq 10\mu m$ [7]) than EASLMs and so greater diffraction angles are possible. Reduction optics are required to image from the LCTV (pixel size $\simeq 80\mu m$) to the OASLM.

III. SYSTEM SCALABILITY

The scalability of this system is determined by various constraints. These include the pitch of the LED/detector array x, the size p of a single hologram pixel at the OASLM, the interconnection neighbourhood size (convolution kernel support) m, the lens focal length f and the hologram width l.

A. Neighbourhood size

The spatial separation of each hologram diffraction order at the exit plane must equal the LED/detector array pitch x. The width d of a single hologram period is therefore given by $d = f\lambda/x$ where λ is the emission wavelength of the LEDs (here assumed to be 660 nm). A single period of the hologram is represented by an $N \times N$ array of pixels, each with a transmittance value of ± 1 [8]. If the neighbourhood size is m then the hologram must diffract light into $m \times m$ diffraction orders. The hologram has only a finite number (N^2) of pixels and so the required diffraction pattern can be obtained with only a limited accuracy. It has been found [9] that m and N are related by $N^2 = m^2/4\epsilon$ where ϵ is the expected deviation of each of the m^2 diffraction orders from its required value. For a binary hologram the diffraction pattern will be symmetric about the origin. If an arbitrary interconnection pattern is required then an off-axis hologram must be used and the number of pixels required to maintain the same coding error ϵ will be doubled. From these considerations we find that the neighbourhood size m for an off-axis hologram is given by $m = f\lambda\sqrt{2\epsilon}/px$ This is an important result in that it specifies the maximum neighbourhood size in terms of the parameters of the optical system and the hologram design error ϵ.

B. Power and speed considerations

The modulation rate of the LEDs will be limited by the received power at the photodetector array. The most greatest insertion loss arises from the fact that LEDs are almost perfect Lambertian sources. Then insertion loss η is approximately given by $\eta \simeq 1 - \cos^2(l/2f)$ where l is the hologram aperture. If $f = 100$ mm and $l = 50$ mm then this component is -12dB. Other sources of loss include hologram diffraction efficiency, intrinsic fan-out loss and excess losses from absorption and reflection.

C. Optical Resolution

The resolution of the system must be sufficiently high that each optical channel can be defined without excessive crosstalk due to diffraction. The smallest aperture in the system will be the hologram aperture (width l). The output image g will therefore be convolved with a sinc diffraction term which has a half-width equal to $f\lambda/l$. If the width of the central lobe of the sinc diffraction envelope is to be less than one-quarter of the LED/detector pitch x then we require $l \geq 8f\lambda/x$. Therefore the LCTV must have $8m/\sqrt{2\epsilon}$ pixels along each side.

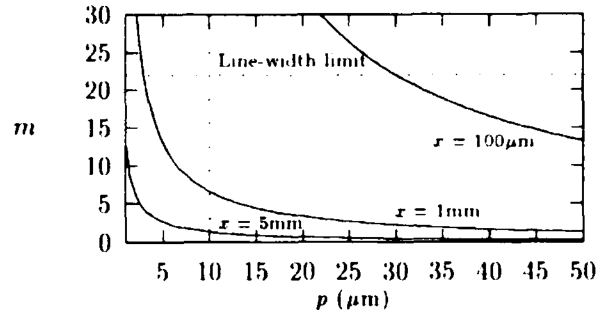
D. Source line width

All previous calculations have assumed that a monochromatic source is used. LEDs have a relatively broad emission spectrum and so the hologram diffraction orders will be broadened. This broadening will be most severe for diffracted orders furthest from the zero-order. For an on-axis hologram the highest order will be located a distance $mx/2$ from the zero-order. If the source line-width is $\Delta\lambda$ then the maximum displacement Δx is given by $\Delta x = mf\Delta\lambda/4d$. If we require that this is less than or equal to one quarter of the LED/detector pitch x then this results in the requirement that $\frac{\Delta\lambda}{\lambda} \leq \frac{1}{m}$

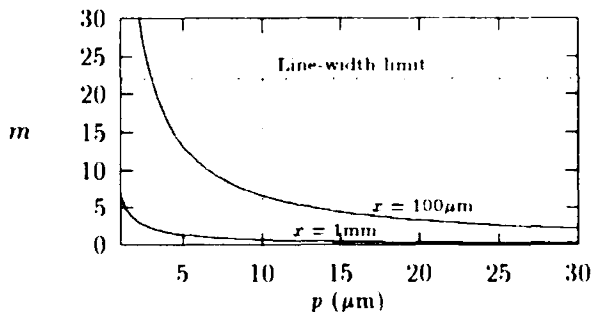
E. System performance limits

The consequences of these factors are illustrated in Fig.3. These graphs show the neighbourhood size as a function of the FELC resolution limit p for a variety of LED/detector pitches x. In Fig.3(a) it is assumed that a 1 m focal length lens is used while in Fig.3(b) a more practical 100 mm focal length lens is assumed. The hologram design error ϵ is specified at 0.5% in both cases. The dotted vertical line at 10 μm represents the maximum resolution of currently available FELC OASLMs and the dotted horizontal line at $m = 22$ represents the maximum neighbourhood size for a LED line-width of 20 nm. It can be seen that a neighbourhood size of $m = 7$ is the maximum which is possible for the proposed system with f =1 m and x =1 mm. In the future fabrication of PE arrays with a pitch of 100 μm should be possible and so the same neighbourhood size may be obtained within a more compact (f = 100 mm) system.

Resolution considerations require that the LCTV has at least 560 x 560 pixels. This is approaching the limit for currently available LCTVs. The total optical losses within such a system will be very high due to the small hologram aperture. An integrated system with $m = 7$, f = 100 mm and x = 100 μm would have a total insertion loss of approximately -57 dB. If high brightness (5mW) LEDs are used this will result in a received power of - 50 dB. This will allow a frame rate of 10^7 frames/second with a 10^{-9} bit error rate and so these losses will not prevent high speed processing operations.



(a)



(b)

Figure 3: Neighbourhood size m as a function of OASLM resolution p for several values of detector pitch x with (a) f = 1 m and (b) f = 100 mm.

IV. Preliminary experiments

In order to investigate the practicality of this system several preliminary experiments have been made with a fixed CGH. The experimental system is shown in Fig.1. The input was provided by a 2 x 2 array of LEDs (pitch 5mm) and a CCD camera was used to simulate the detector plane. A shift-invariant interconnection operation was performed by a non-seperable binary CGH designed by simulated annealing [8] and fabricated as a photoresist on glass device. The hologram was designed to perform a 5 x 5 uniform fanout operation and had a period size of 128 μm and an aperture size of 4 mm. The lens focal lengths used were f_1 = 1 m, f_2 = 100 mm. Fig.4 shows the fanout operation which is performed when a single LED was illuminated. The diffraction efficiency of the hologram was found to be 40% and the fanout non-uniformity was ±10%. The bright zero-order was due to an incorrect CGH phase depth. Three LEDs (shown in Fig.5(a)) were then used to provide the input to the convolution system. Due to alignment errors there was a large variation in the power received at the exit plane from each LED. The insertion loss due to the small CGH aperture was found to be -44 dB. Fig.5 also shows the convolution output together with the intensity profile of the fifth row. The left-most peak has a contribution from only one LED, the rightmost has contributions from 2 LEDs and the central 4 peaks represent the sum of all 3 LEDs. It was found that after adjusting

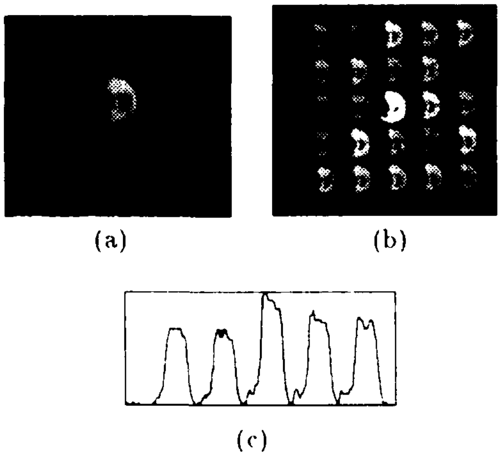


Figure 4: (a) Single LED image, (b) after multiple imaging with CGH, (c) intensity profile.

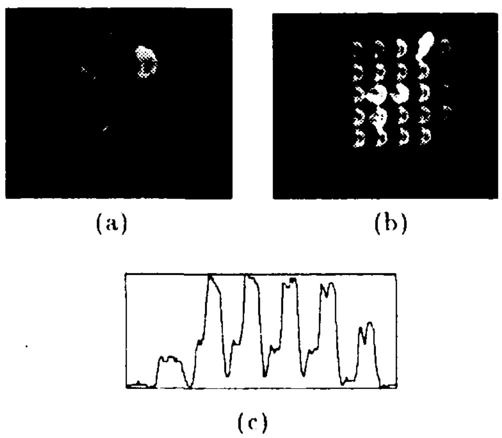


Figure 5: (a) 3 LED image, (b) after multiple imaging with CGH, (c) intensity profile through 5th row.

for the non-uniform power contribution from each LED the error in the convolution operation was less than 5%.

V. Conclusions

A reconfigurable shift-invariant optical interconnection scheme for cellular and early vision processing has been presented. This system will interconnect an integrated array of high-speed electronic processing elements with optical input and output. The interconnection operation is performed by the use of reconfigurable Fourier plane diffractive elements. Interconnection neighbourhood sizes of 7 x 7 are possible with currently available components and future improvements in OASLM resolution will enable neighbourhood sizes of more than 20 x 20 to be obtained. The system frame rate is predicted to be 10^7 frames/second (limited by received power). Reconfiguration time will be determined by the FELC/LCTV update time and may be limited to video frame rates.

Experimental investigations have been made with a LED array input and an $m = 5$ interconnection neighbourhood size provided by a fixed CGH. Although received power was low due to the small LED aperture the interconnection error was less than 5%. Further work is now required to implement the reconfigurable system.

Acknowledgements

The hologram used here was designed and fabricated in the Department of Physics, King's College London, UK.

References

1. C Mead. *Analog VLSI and Neural Systems.* Addison-Wesley, Massachussets, 1989.
2. M Ishikawa, A Morita, N Takayanagi, 'High speed vision system using massively parallel processing' *Proc. Intl. Conf. on Intelligent Robots and Systems* pp 373-377, 1992.
3. B K Jenkins, P Chaval, R Forchheimer, A A Sawchuck, T C Strand, 'Architectural implications of a digital optical processor', *Appl. Opt.* **23** (19) pp 3465-3474, 1984.
4. Y Takaki, H Ohzu, 'Holographic function/interconnection module', *Appl. Opt.* **30** (20) pp 2825-2832, 1991.
5. A G Kirk, G D Kendall, M-Y Chan, T J Hall, 'Holographic Interconnections within cellular optical processing systems', *to appear in Optical Computing and Processing*, 1992.
6. D C O'Brien, W A Crossland, R J Mears, ' A holographically routed optical crossbar: theory and simulation', *Optical Computing and Processing* 1 (3) , pp 233-244, 1991.
7. N Kato, R Sekura, J Yamanaka, T Ebihara, S Yamamoto, 'Characteristics of a ferroelectric liquid crystal spatial light modulator with a dielectric mirror', *Proc SPIE* **1455**, pp 190-205, 1991.
8. M A Seldowitz, JP Allebach, DW Sweeny, 'A synthesis of digital holograms by direct binary search', *Appl Opt* **26** (14) p. 2789, 1986.
9. A G Kirk and T J Hall, 'Design of computer generated holograms by simulated annealing: coding density and reconstruction error', to appear in *Opt. Commun.*, 1992.

Optical Interconnections: 2

OFA 8:30am-10:00am
Grand Ballroom West

Phillip J. Anthony, *Presider*
AT&T Bell Laboratories

GaAs Heteroepitaxial Growth on Submicron CMOS Silicon Substrates

S.K. Tewksbury, L.A. Hornak, H. Noriman, S. McGinnis
West Virginia University
Engineering Sciences Bldg. 827
P.O. Box 6101
Morgantown WV 26506-6101

Preliminary results of our study of the CMOS compatibility of GaAs heteroepitaxial growth on wafers containing twin-tub V CMOS devices (0.9 micron drawn features) are reviewed.

Four- foci slanted axis Fresnel lens for planar optical perfect shuffle

C.D.Carey, D.R. Selviah, J.E.Midwinter,
Department of Electronic and Electrical Engineering, University College London, Torrington Place, London
WC1E 7JE, UK (Tel) +44 71 387 7050 (Fax) +44 71 387 4350 (e-mail) d.selviah@ uk.ac.ucl.eleceng

S.H.Song and E.H.Lee
Electronics and Telecommunications Research Institute, Research Department, P.O.Box 8, Daedog Science Town, Daejon 350-606, Rep. of Korea(Tel) +82-42-860-6033 (Fax) +82-42-860-5033 (e-mail) shsong@ard.etri.re.kr

1. INTRODUCTION

Planar optical technology offers a promising compact and robust approach for performing complex optical interconnects for silicon VLSI and WSI circuits. A transparent parallel sided medium is mounted over the circuit using solder bump mounts [1]. Arrays of reflection modulators or surface emitting laser diodes are imaged obliquely via multiple reflections within the interconnect substrate (Fig 1) and via reflection holographic lenslets, routing elements, microprisms and microlenses to arrays of detectors. Crossover and perfect shuffle interconnect patterns are important for switching systems and already planar optical crossover interconnects using microprisms have been demonstrated [2]. In this paper we describe the design and demonstration of holographic elements for performing two dimensional folded perfect shuffles in a planar optical configuration. Optical demonstrations of the perfect shuffle have used bulk lenses and prisms [3] but of particular relevance here are the demonstrations using arrays of 2x2 lenses [4] each of which both magnifies the original image and creates 4 overlapping copies of it. Such arrays of lenslets can be recorded holographically and will perform the same operation [5]. To allow mass fabrication, however, the hologram needs to be easily designed, fabricated and copied so binary, surface relief, computer generated holograms are desirable. For a compact system the focal lengths of the lenses need to be ~1cm which means that the holograms have to be designed using Fresnel diffraction theory. For an input image consisting of small pixels only a small lateral shift is required so the 4 foci of the lenslets need to be spaced about 60-250μm apart. One approach is to split the hologram plane into four and to place in each quadrant only a part of the hologram required for a particular focus [6]. This approach is sensitive to the illumination intensity distribution and can be expected to yield asymmetric foci. A second approach is to calculate the interference pattern due to all four foci and the illuminating beam over the whole hologram plane and to threshold it to make it binary. This approach requires larger memory capacity but has the advantage that regular smaller foci can be achieved and the design is robust due to its redundancy [6]. Design equations have been derived [7] and normal axis, four-foci, reflecting, Fresnel lenses of this type have been demonstrated [8]. Slanted axis 1:4 fanout reflection holograms of this type have also been described but they were not configured for a planar optical implementation and their aim was simply for clock distribution rather than imaging [6].

2. FOUR-FOCI SLANTED AXIS HOLOGRAM DESIGN

The hologram was designed to be a binary, reflective, amplitude structure operating at 633nm wavelength in a parallel sided quartz glass substrate of 6 mm thickness. In the metallised regions the light was to be reflected and in the transparent regions transmitted. The refractive index of quartz glass at this wavelength is 1.457 corresponding to a critical angle of 43.2° so to ensure transmission a smaller angle of incidence must be chosen and 30° was chosen. It was noted that for such angles light polarised parallel to the surface gave slightly better transmittance. The hologram was designed (Fig. 1) so that the focal length corresponded to two oblique traverses across the thickness of the substrate, f=12mm. The 2f imaging distance corresponds to four traverses of the substrate and was chosen to place the input array, hologram and output array on the same plane for ease of alignment. The interference pattern was calculated between one 30° angled collimated beam from the source and four beams converging to point foci at the required angles for a planar optical configuration. This was done by dividing the hologram plane into 1μm pixels and in each pixel the intensity was calculated by summing the 5 beams taking account of their phases. The four foci were chosen to be in phase on the focal plane and, for the first design, to lie in a square configuration of side 250μm and, for the second design, at the vertices of a rhombus of major diagonal $350\sqrt{2}$ μm and minor diagonal $150\sqrt{2}$ μm. The point sources were assumed to be isotropic and the inverse square dependence of amplitude on distance propagated was neglected. The following equation was used to calculate the intensity of illumination on the hologram plane at pixel location (x,y) neglecting any constant illumination term

$$\Delta I(x,y) = \sum_{i=1}^{4} \cos(\phi_i - \phi_r)$$

where ϕ_i is the phase at the hologram plane due to the ith focus and ϕ_r is the phase due to the oblique collimated beam. When ΔI was greater than zero the pixel was assigned to be reflective otherwise it was assigned to be transmissive. This calculation was performed over a square of side 1mm which limited the memory requirement to 3.7MBytes for a GDSII format file. After fabrication of a chrome on glass mask by electron beam the pattern was copied lithographically to give a reflective aluminium pattern on the quartz substrate. The patterns for the two holographic lens designs are shown in figures 3(a) and (b) for the square and rhombus focal array.

3. PLANAR OPTICAL PERFECT SHUFFLE

The principle of the formation of the perfect shuffle is shown in figure 2. The hologram can be considered to consist of a 2x2 multiplexed array of lenslets. Each lenslet images the input object array to the output but shifted by a distance equal to the spacing of the focal points. First consider the square focal point array. If a square input array of side equal to the focal point spacing is used as input then on the output plane one obtains four copies side by side in a square configuration; each quadrant contains one copy (Fig 2). If the situation were reversed and an input array consisting of four quadrants were placed at the input, by reciprocity all four quadrants would be superimposed at the output. Now, if each input quadrant contains a 2x2 sub-array of illuminated pixels separated by one or more pixel widths, each of side less than or equal to one quarter of the quadrant side, and if in each quadrant the sub-array were shifted towards a different corner of the quadrant one could obtain the pattern shown in figure 4(a) for one choice of shifts. Clearly when each of the quadrants are superimposed the 2x2 pixel arrays are interleaved performing a folded perfect shuffle. Another choice of shifts would move all the 2x2 arrays together so that the central 4 pixels move together while yet a further choice would move them all apart. In any of these configurations the input array is not regular. This was solved by arranging for the foci to be spaced so that the hologram performs the shifts itself. For the choice of shifts shown in figure 4(a) two opposite foci are moved together by $100\sqrt{2}\,\mu m$ and the other two are moved further apart by $100\sqrt{2}\,\mu m$ to give a rhombus focal array. The experimentally obtained output for the square focal array designs (Fig 4(b)) shows a perfect shuffled output in the central region of four by four pixels. This was verified by covering the input pixels in turn and noting the missing pixels in the output pattern. In the experimental results shown incoherent white light illumination was used which has the advantage that it gives reduced speckle at the output and alignment is not so critical as the wavelength spread results in a range of focal lengths for the lens.

4. CONCLUSIONS

We have described the design of a four-foci, square focal array, 30^{o} slanted axis, binary, reflective, amplitude, computer generated holographic, Fresnel lens and have demonstrated its use in a planar optic imaging configuration for performing a perfect shuffle using incoherent white light illumination. The design of a second rhombus focal array, four-foci, Fresnel lens is described and shown for perfect shuffling a regular input array of pixels in the planar optic configuration. Higher efficiency developments of the latter design for infra red wavelengths will be valuable for perfect shuffle optical interconnects between regular arrays of modulators and detectors on VLSI and WSI circuits for high speed switching systems. This work was supported by the UK SERC via the UCL Optoelectronic Rolling Grant and by the Korean Government. The authors thank Dominic Godwin for helpful discussions.

REFERENCES

1 J. Jahns, Y. H. Lee, C. A. Burrus Jr and J. L Jewell: "Optical interconnects using top-surface-emitting microlasers and planar optics", Applied Optics, **31**(5), 592-597, (1992)

2. S. H. Song , S. H. Lee, C. D. Carey, D. R. Selviah and J.E. Midwinter, "Planar optical implementation of crossover interconnects", Optics Letters, **17**(18), 1253-1255, (1992)

3. A.W.Lohmann, "What classical optics can do for the digital optical computer", Appl. Opt., **25**(10), 1543-1549 (1986)

4. M.W.Haney and J.J.Levy, " Optically efficient free-space folded perfect shuffle network", Appl. Opt., **30**(20), 2833-2840 (1991)

5 N.Davidson, A.A.Friesen and E.Hasman, "On the limits of optical interconnects", Appl. Opt., **31**(26), 5426-5430, (1992)

6 H.Zarschizky, H.Karstensen, A. Staudt and Ek. Klement: "Clock distribution using a synthetic HOE with multiple fan-out at IR-wavelength", Proc. SPIE Int. Soc. Opt. Eng. ,1281, 103-112 (1990)

7 M.Kato and K.Sakuda, "Computer-generated holograms: application to intensity variable and wavelength demultiplexing holograms" Appl. Opt., **31**(5), 630-635 (1992)

8 C.D.Carey, D. R. Selviah, S. K. Lee, S. H. Song and J. E. Midwinter, "Computer generated hologram etched in GaAs for optical interconnect of VLSI circuits", Elect. Lett., **28**(22), 2082-2084 (1992)

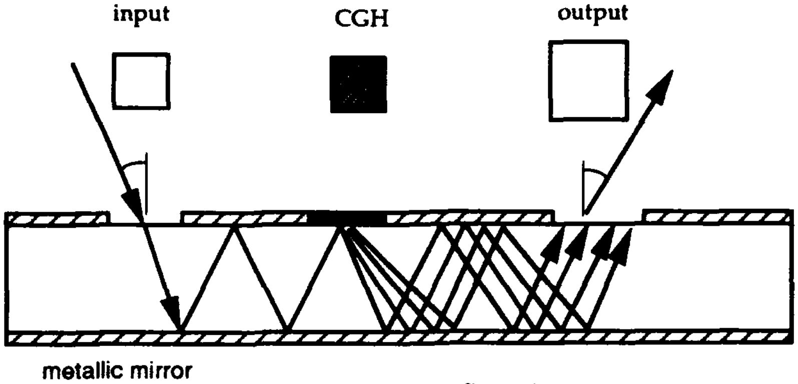


Fig.1 Schematic of planar optical perfect shuffle configuration

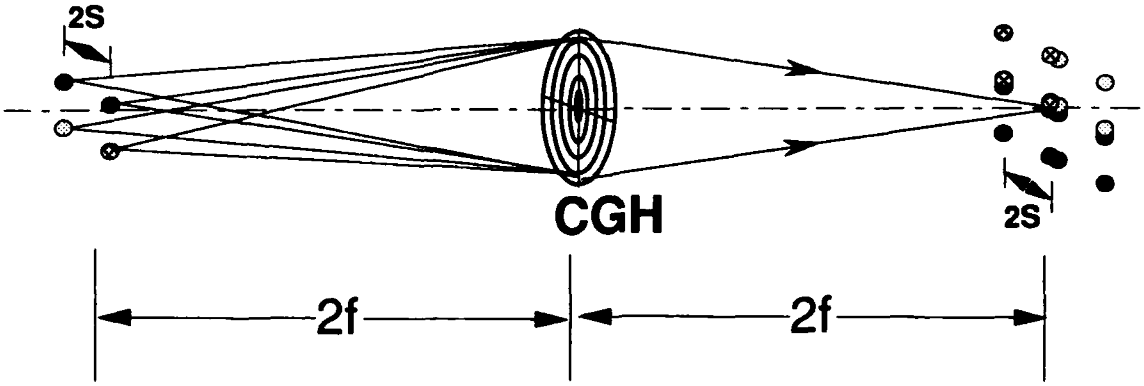
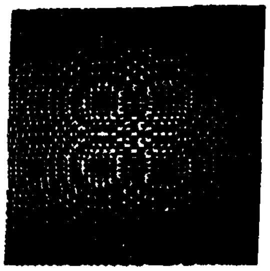
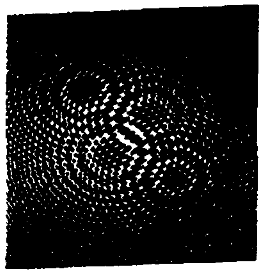


Fig.2 Perfect shuffle operation using 2X2 array of multiplexed lenslets



(a)



(b)

Fig.3 Four-foci, 30 degree slanted axis, computer generated holographic Fresnel lens, f=12mm, (a) square focal array ,(b) rhombus focal array

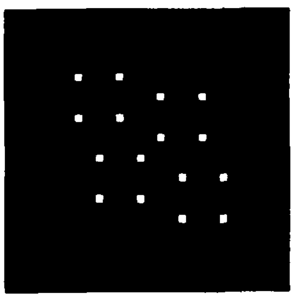


Fig.4(a) 4X4 Input pattern of four shifted quadrants

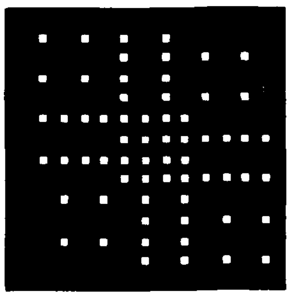


Fig.4(b) Theoretically predicted output pattern using square focal array lens

Fig.4(c) Experimental perfect shuffle pattern (4X4 array in centre)

Experimental implementation of a 25 channel free-space optical switching architecture

Andrew Kirk* Stuart Jamieson Hussain Imam Trevor Hall

King's College London, Strand, London WC2R 2LS Tel. +44 71 836 5454 ext. 3718

* now at University of Tokyo, Bunkyo-ku, Tokyo 113 Japan Tel. +81 3 3812 2111 ext. 6902

I. INTRODUCTION

There are many areas within optical telecommunications and optoelectronic information processing in which an optical space-switch is required to reconfigurably interconnect two arrays of optoelectronic devices. The authors have previously described a candidate optical space-switch which uses holographic multiple imaging together with ferroelectric liquid crystal (FELC) spatial light modulators (SLMs) [1,2]. This architecture is attractive in that it will scale to allow the interconnection of more than 200 channels, it is compact and may be implemented with currently available components. In this paper we describe the implementation of a 25 channel version of this system. Experimental results are presented together with a consideration of the implications which these have for large-scale implementations of this system.

The system described here is a development of the well-known vector-matrix multiplier crossbar switch. This was originally developed by Goodman et al in 1978 [3] and has since been used widely in many optoelectronic information processing systems [4]. The scalability of this system can be greatly improved by redesigning it as a two-dimensional matrix-matrix multiplier [2].

II. MATRIX-MATRIX MULTIPLIER

A schematic diagram of this system is shown in Fig.1 [2]. The switch interconnects N^2 optical channels which are arranged as an $N \times N$ array. Light enters the system as a polarised Gaussian beam which illuminates the Fourier plane array generator (fan-out hologram) H_1. This creates an array of $N \times N$ Gaussian beamlets, each of which illuminates a single pixel of the electrically addressed ferroelectric liquid crystal (FELC) spatial light modulator SLM1. This SLM provides the input to the switch. After passing through an analyser (not shown) the modulated beamlet array is recollimated and illuminates a second Fourier plane array generator H_2. This is also

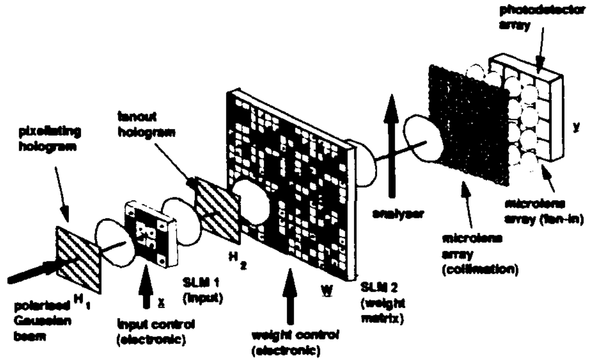


Figure 1: Schematic diagram of the matrix-matrix multiplier.

an $N \times N$ fan-out hologram but has a spatial frequency which is N times that of H_1. This causes the original $N \times N$ beamlet array to be multiply imaged $N \times N$ times at the second FELC SLM (SLM2) [5,6]. This provides the interconnection weight matrix and has $N^2 \times N^2$ pixels. After the polarisation state of each beamlet has been modulated by SLM2 the array is recollimated and passes through a second analyser before being imaged once more at the microlens array. Each beamlet is individually collimated by a single microlens and then focused onto a photodetector by an array of $N \times N$ fan-in lenses. The matrix-matrix multiplication output is obtained at the $N \times N$ photodetector array. The two fan-out holograms in this system thus act as cascaded array generators [7].

Two major advantages are gained by redesigning the free-space optical crossbar switch in this way. By using holographic multiple imaging the fan-out operation is spread over two dimensions and so the degree of fan-out required for a single dimension is reduced. Thus a 256 channel crossbar requires a 1 to 256 fan-out operation for the vector-matrix multiplier but only a 1 to 16 $\times$ 16 fan-out operation in the case of the vector-matrix multiplier. Holographic fan-out of up to 128 $\times$ 128 has been reported [8] and so such an operation is feasible. The second advantage is that optical signals are propagated through the system as an array of Gaussian beamlets. If the SLM pixels are sufficiently large then a single beamlet will pass through the pixel with negligible uncontrolled diffraction from the edges of the pixel. It has been shown that diffraction of this type can seriously reduce the scalability of optical interconnection systems [9]. These factors allow a very small SLM pixel pitch to be used and so in principle a very compact system may be constructed. Studies of this system suggest that optical constraints should allow a 256 channel system to be implemented within a cylinder with a diameter of 1.7 mm and a length of 26 mm [2]. Practical considerations of SLM fabrication and power dissipation may require a system that is somewhat larger than this however.

III. Experimental Implementation

A. Hologram fabrication and evaluation

A 25 channel version of the system described above requires two 5 × 5 fan-out holograms. A 5 × 5 non-separable binary fan-out computer generated hologram (CGH) was designed by use of the simulated annealing algorithm [10]. This hologram had a simulated diffraction efficiency of 70.0% and an intensity non-uniformity of 0.3%. The design was then fabricated directly as a resist-on-glass pattern by use of a direct-write electron-beam plotter. Two holograms with period sizes of 320 μm (H_1) and 64 μm (H_2) were fabricated. The width L of both holograms was 2mm. The experimental fan-out non-uniformity of the two holograms was found to be 8.5% with a diffraction efficiency of 70%. The degradation in image resolution caused by the holograms was determined by multiply imaging a resolution test chart. It was found that a 20% loss in resolution occurred when compared with the resolution obtained by imaging with an aperture of the same dimensions.

B. Experimental system

The experimental matrix-matrix interconnect system which was constructed is shown in Fig.2. This is very similar to the system shown in Fig.1. In the experimental system the complete fan-in stage was not implemented. Light from a single sub-array at SLM2 was collected into a photodetector in order to obtain the matrix-matrix multiplication result for a single output channel. The full switched image at SLM2 was also observed at the CCD camera. All the lenses used were f/3.3 achromatic plano-convex with a focal length of 100 mm. The SLMs used were

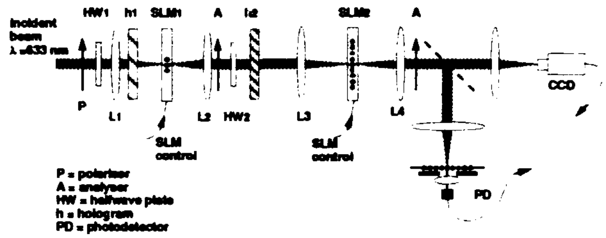


Figure 2: Experimental system.

128 × 128 electrically addressed FELC SLMs manufactured by STC Ltd (now BNR). These had a pixel size of 165 μm. Two microcomputers were used to control the SLMs.

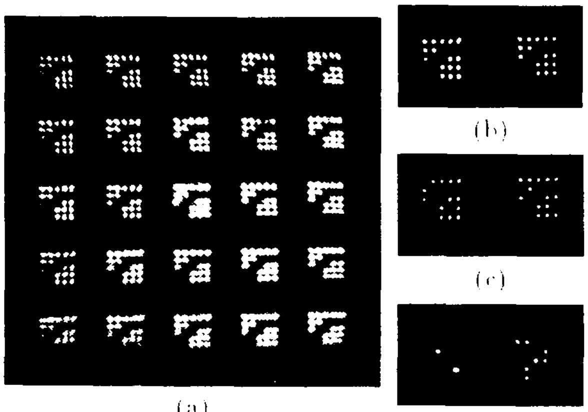
C. Contrast ratio and insertion loss

The effect which the holograms had upon the system contrast ratio was measured. It was found that when the SLMs and polarisers were absent the holograms and lenses caused an 11 dB reduction in the extinction ratio of the system. The contrast ratio of the two SLMs was also measured. When all pixels in SLM1 were in the 'on' state the contrast ratio for a single pixel at SLM2 was found to vary between 35 and 70. When all pixels at SLM2 were switched to 'on' the contrast ratio at SLM1 was found to be approximately 35. The specified contrast ratio for the SLMs was greater than 150:1. In order to achieve this however it is necessary to synchronise the optical signal observation with the 'pause' period of the SLM addressing signal. It was not possible to perform this operation simultaneously with both SLMs. The contrast ratio was also found to depend critically on the alignment of both SLMs with the cascaded beamlet array.

The excess insertion loss of the system was found to be -23 dB. All the glass surfaces were uncoated and so would have contributed a predicted excess insertion loss of 4.3 dB. The predicted excess insertion loss of the holograms and of the SLMs and the polarisers was estimated at 15dB. Therefore 3.7 dB excess insertion loss was unaccounted for. One possible unmeasured source of loss is the scattering of light by the holograms into unwanted polarisations.

D. Switching operation

Fig.3(a) shows the multiple imaging of a 5 × 5 input pattern presented at SLM1 when all pixels at SLM2 are set to 'on'. By adjusting the state of the pixels at SLM2 different channels can then be switched on



(a) (b) (c) (d)

Figure 3: System output observed at CCD: (a) multiple imaging of pattern presented at SLM1; (b), (c), (d) – details of two sub arrays as SLM2 pixels are switched.

and off. Details of two sub arrays as different pixels in SLM2 are switched are shown in Fig.3 (b), (c) and (d).

E. Signal to noise ratio

When the crossbar switch is used within a telecommunications system the signal in one input channel must be discerned above the rejected noise from all remaining $(N^2 - 1)$ channels. The worst case signal to noise ratio (SNR) was measured for the 25 channel system. A single pixel at SLM1 was switched on together with the corresponding pixel in one output subarray at SLM2 and the transmitted signal strength from the subarray was measured. The pixel at SLM1 was then switched off and all the remaining 24 pixels at SLM1 were switched on and the noise strength was measured. The output of the system is shown as an oscilloscope trace in Fig.4 where the signal level is the upper trace and the noise is the middle trace. The best measured values resulted in a worst case SNR of 6dB. The noise was thought to arise mainly from the finite contrast ratio of SLM2. This SNR level was however sufficient for the 25 channel system to function as a crossbar switch.

F. Switch linearity

In a matrix matrix multiplier the output is given by the weighted sum of the input channels and so the switch linearity is also an important parameter. This was assessed by switching on an increasing number of pixels at SLM2 and measuring the received signal strength. The switch response is shown in Fig.5. It can be seen that when a small number ($\simeq 5$) of pixels were selected at SLM2 the response was linear but when more than 20 pixels were selected the deviation became greater than 1.0. This is thought to have been due to the non uniform fan out of the hologram H_1 (8% non uniformity) and to beamlet misalignment at SLM2 which caused a loss of contrast ratio.

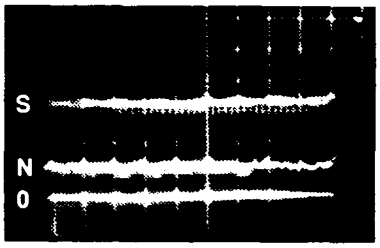


Figure 4: Worst case signal to noise ratio. S – signal, N – rejected noise, 0 – zero level.

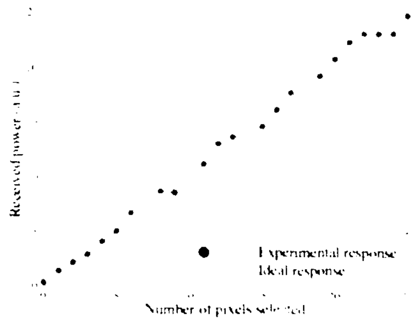


Figure 5: Linearity of response.

IV. Discussion

The most critical factor in this system is the alignment of the various components. When the beamlet array at SLM2 was not precisely aligned with the pixels the contrast ratio was greatly reduced. It was not possible to precisely align all 25 × 25 beamlets at SLM2 with the pixels because of optical aberrations and because of a mismatch between the fan out beamlet pitch and the pixel pitch. With the use of high quality lenses and precision fabrication techniques it should be possible to greatly eliminate these problems. In order to achieve the degree of alignment required however various new techniques must be investigated such as the use of 'lock and key' alignment marks.

The use of anti-reflection coatings will greatly improve the performance of this system. In the present system a significant proportion of the incident light is lost due to unwanted reflections. This scattered light will reduce the SNR of the system. The use of antireflection coatings should also improve the uniformity of the holograms. CGH non-uniformities of less than 1% have been reported when anti-reflection coatings are used [11]. The possible depolarisation effects of the holograms should also be investigated. The scalability of the system is limited mainly by the contrast ratio of the SLMs [2]. It is therefore important that the polarisation state of the light in the system is maintained as accurately as possible. A significant improvement in the contrast ratio of this system would also be obtained if the high-contrast read period of the both SLMs was synchronised.

V. Conclusions

The vector-matrix multiplier has been redesigned as a matrix-matrix multipier with two-dimensional fanout obtained via holographic multiple imaging. A 25 channel version of this system has been implemented experimentally with a single operational output channel. By routing a beamlet array through two successive electrically addressed SLM pixel arrays these experiments have demonstrated the viability of the concept of the use of Gaussian beamlet arrays as free-space optical channels. The experimental system suffered from a high excess insertion loss (-23 dB) due to the large number of optical components in the system. In principle however this can be greatly reduced by the use of anti-reflection coatings. The experimental system was found to have a 6 dB worst-case signal to noise ratio. It was found that the most important factor within this system is the alignment of the beamlet array with the pixel array at the SLMs. In the experimental system it was not possible to achieve a good alignment of all 25×25 beamlets with the pixels of SLM2 simultaneously. Further investigation must therefore be made in this area in order to determine alignment tolerances and possible alignment techniques. Additional work is also required to implement the complete matrix-matrix multiplier design with all output channels operational. This will require the use of microlens fan-in optics together with a photodetector array.

Acknowledgements

This work was partially funded by the UK SERC/DTI LINK Optoelectronics project 'Holographic interconnects for parallel processing systems' (HIPPS). The authors are grateful to British Aerospace plc and British Telecom plc for sponsorship which supports their work. The authors would like to thank Bill Crossland of BNR Europe for assistance with the SLMs and for much useful discussion.

References

1. A G Kirk, H Imam, K Bird and T J Hall, 'The design and fabrication of computer generated holographic fanout elements for a matrix-matrix interconnection scheme', Proc. SPIE Vol. **1574**, 1991.
2. A G Kirk, G D Kendall, M-Y Chan, T J Hall, 'Holographic Interconnections within cellular optical processing systems',to appear in *Optical Computing and Processing*, 1992.
3. J W Goodman, A R Dias, L M Woody, 'Fully parallel high-speed incoherent optical method for performing discrete Fourier transforms', **2**, (1), p 1-3, 1978.
4. K M Johnson and G Moddel, 'Motivations for using ferroelectric liquid crystal spatial light modulators in neurocomputing', *Applied Optics* **28** (22), pp 4888-4899, (1989) and references therein.
5. G Groh, 'Multiple imaging by means of holograms of correlated objects', *Appl. Opt.* **7**, pp 967-969, 1969.
6. K J Weible, G Pedrini, W Xue, R Thalman, 'Optical interconnection of a neural network associative memory using diffraction gratings', *Japanese J. Appl. Physics* **29** (7) pp L1301-L1303, 1990.
7. F B McCormick, 'Generation of large spot arrays from a single laser beam by multiple imaging by means of binary phase gratings', *Opt. Eng.* **28** (4) 299-304, 1989.
8. M R Taghizadeh, J Turunen, B Robertson, 'Passive Optical Array Generators', OSA 1991 Technical Digest Series, Vol 6, p 148, 1991.
9. J Shamir, H J Caulfield, R B Johnson, 'Massive holographic interconnection networks and their limitations', *Appl. Opt.*, **28** (2), pp 311-324, 1989.
10. M A Seldowitz, JP Allebach, DW Sweeny, 'A synthesis of digital holograms by direct binary search', *Appl Opt* **26** (14) p. 2789, 1986.
11. M.P.Dames, R.J.Dowling, P.McKee, D.Wood, 'Design and fabrication of efficient optical elements to generate intensity weighted spot arrays', *Applied Optics* **30** (19), pp 2685-2691, 1991.

Bandwidth as a Virtual Resource
In Reconfigurable Optical Interconnections

Donald M. Chiarulli, Steven P. Levitan, Rami G. Melhem, Chunming Qiao
Dept. of Computer Science
University of Pittsburgh
Pittsburgh, PA

This work is focused on reconfigurable optoelectronic interconnection networks: networks constructed of optical waveguides in which messages are switched or routed by means of optoelectronic devices [Goo89]. The dichotomy between the bandwidth of the optical channels which carry information through these networks and the performance of the electronic controllers and decoders which determine the routing and destination of those messages is a significant bottleneck. In these networks, addressing and control are implemented jointly as message routing. Unlike addressing, routing considers not just the destination of a message, but also its path, and the network resources needed for that path. Any routing strategy represents a tradeoff between explicit addressing and global control. In most systems, explicit addressing dominates this tradeoff. In other words, source nodes drive the control hardware which arbitrates resources to create the message path.

In this paper we present an alternative in which the control system dynamically allocates resources based on global knowledge of the message traffic. Thus, rather than transmitters presenting addresses to the network, the network establishes a set of paths and presents these to the transmitters and receivers. Since most interconnection networks cannot provide all possible paths simultaneously, the issue in these routing strategies is to allocate the network resources such that they consistently meet the needs of the current message traffic. If a sequence of network configurations exsist which meet the needs of the current message traffic, then control operations will only be required to tranform this sequence to track the changes in this traffic. The locality which is inherent in the message traffic, suggests that such changes will occur slowly relative to the total volume of message traffic. Thus the performance of the control system can be decoupled from the message throughput of the network.

Network Model

The general structure of a communication network based on this paradigm is shown in figure 1. Let INET in this figure be an $n \times n$ interconnection network connecting a set I of n input ports to a set O of n output ports, and let $p_{io} = (i, o) \in I \times O$ denote the path between a specific input port $i \in I$ and a specific output port $o \in O$. We assume that the INET may establish any of the possible $N = n^2$ paths, but that it cannot establish them simultaneously. Thus, INET may be a bus, a multistage interconnection network (MIN), a wavelength division switched (WDS) star, or any other type of interconnection network. Within an INET, we define a mapping, M, to be a set of paths that can be established at the same time without conflicts in the INET. For each mapping M there is a corresponding state S which represents the configuration of the network (i.e., switch settings, detector tunings, etc.) corresponding to that mapping.

Since the establishment of two paths at the same time in an INET may cause conflicts, not every set of paths is a mapping. We refer to the establishment of all the paths in a mapping as the *realization* of the mapping. Given a set of paths $P \subseteq I \times O$, it may not be possible to realize all paths in P at the same time without conflicts. However, P can be partitioned into several mappings, $P = M_1 \bigcup M_2 \bigcup ... \bigcup M_t$. Each mapping M_i, $i = 1, 2, ..., t$, may be realized in sequence. Given that each mapping has a corresponding state, the set of paths P may be implemented as an

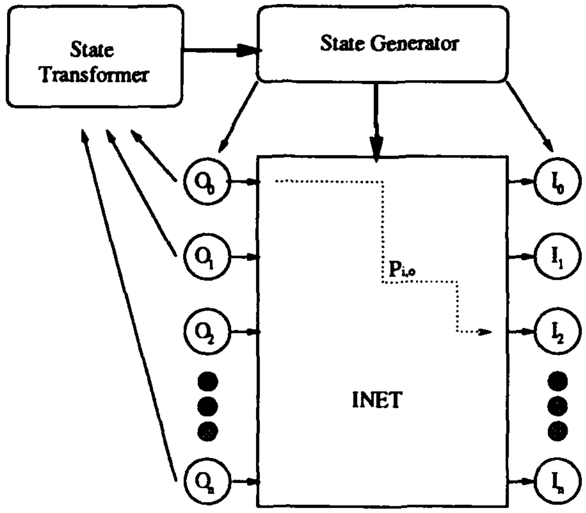


Figure 1: **General Interconnection Block**

ordered sequence of states, $[S_1, S_2, ..., S_t]$ where t is the length of the sequence.

Returning now to figure 1, the state generator block is responsible for generating the current state sequence. The control algorithm, which determines the sequence, runs in the state transformer. The current state of the network is also communicated to each of the transmitting nodes. Thus, a transmitting node waits for the network state corresponding to a mapping which contains the required path. When such a state is detected, the node transmits its message. If no such mapping exists within the current state sequence, the control algorithm modifies the state sequence to include a mapping which supports the requested path. In the next section we show that this modification operation can be modelled directly by a page replacement operation in virtual memory system or cache manager.

Routing as selection in a virtual connection space

We draw a direct analogy between the routing problem for reconfigurable multiprocessor interconnection networks and paging in virtual memory systems. The correspondence between the two is summarized in table 1. As shown in table 1, connections in a fully connected network can be viewed as the analog of memory locations in a complete virtual address space. Just as a physical memory supports a subset of the virtual address space, so a switched interconnection network implements a subset of connections from the fully connected network. Physical memory is shared and reused to create the illusion of large virtual memory space. Similarly, a switched network can be reconfigured to emulate the functionality of full interconnection.

The analogy also extends to addressing. The unit resource in a memory system is a single memory location. In a communication network the unit resource is a single connection path. In memories, an $n - bit$ virtual address defines an address space of $2^n = N$ addressable locations. Paging divides this address space into k pages, each of size m locations such that $N = m \times k$. In communication networks, a full interconnection network for n sources and n destinations provides $n \times n = N$ interconnection paths, i.e., N unique source-destination pairs. While an arbitrary switched interconnection network may establish any of these N paths, it is only capable of connecting a subset of m paths simultaneously, giving a particular configuration. In order to enumerate all possible paths, a *sequence* of k different configurations is required, such that $N = m \times k$. Just as at any given time, a subset of the k pages resides in physical memory to satisfy the current set of memory requests, a communication network needs only to sequence through a subset of configurations to satisfy the current requests for paths.

Entity	in Virtual Memory	in Communications Network
Addressing Space	Virtual Address	$n \times n$ connection space
Shared Resource	Physical Memory	Mapping sequence
Sharable Unit	Page	Mapping
Addressable Unit	Memory Location	Path

Table 1: Summary of Virtual Memory vs Switched Interconnection Analogy

Virtual memory systems work because the principle of locality states that if a working set of pages is made available to a group of programs, that set of pages will change slowly over time. By extension, if a sequence of configurations is sufficient to support all of the current traffic in an interconnection network, so will that sequence change slowly over time. Thus, we can use the principle of locality to decouple the performance of control algorithms for interconnection networks from the latency of individual messages. The control algorithm needs only to perform in the time frame of locality changes, not in the time frame of individual message traffic. This is a key concept for high bandwidth optical interconnection networks. Since the routing decisions in these networks are most often made by electronic controllers, the performance of these controllers represents a significant bottleneck. By exploiting locality, message routing can be reduced to a problem of providing a repeated sequence of configurations to the network. Control becomes a problem of transforming that sequence to track the changes in the locality of message traffic.

Examples

Figure 2(a) is a multistage interconnection network (MIN). In this network, a unique path can be established between any input port and any output port. Along each path, there are $\log n$ switches, one at each stage. In order to establish a path, each switch along the path has to be set properly to either a "straight" or a "cross" state. Other switches in the network can be in either state without affecting the path. Therefore, in any mapping there will be exactly n paths. Given a mapping, it is straightforward to find the state S that realizes the mapping. Once the set of mappings is determined, the state generator outputs the sequence of states which establishes the mapping for all current paths. As described earlier, each transmitter monitors the output of the state generator and waits for a mapping which contains the requested path.

Should a path be requested which is not contained in the mapping sequence, the state transformer must replace one of the mappings which a new mapping which contains the required path. In most cases this will require the preemption of paths in the mapping to be replaced. Just as in optimal page replacement, the optimal algorithm for this selection is one which choses the mapping with paths needed most distant in the future. Random, FIFO, LRU, or NUR algorithms can all be applied in this context.

Another example shown in figure 2(b) is a wavelength division multiplexed star network. A variety of interconnection protocols exist for these nets [Dou91]. For this example, we assume the most general case in which any of I transmitters may communicate with any of O receivers on any of W wavelengths. In other words, transmitters and receivers are independently tuned. Since in general $\mid W \mid \leq min(\mid I \mid, \mid O \mid)$, we further assume that any transmitter or receiver may be effectively "turned off" such that it will neither transmit or receive. Under these assumptions, a path consists of a triple $(i, o, w) \in I \times O \times W$ and a mapping consists of any realizable set of paths such that each path communicates on an different wavelength. The state S corresponding

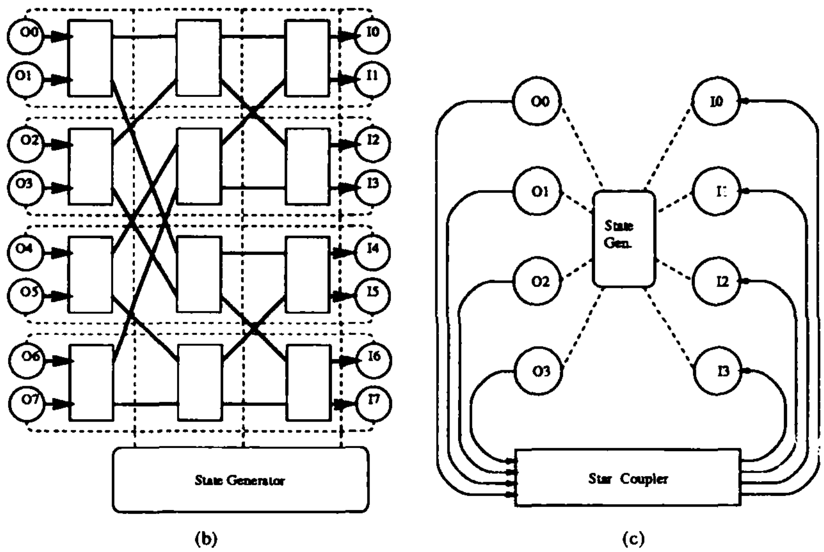


Figure 2: **Two Interconnection Networks**

to each mapping is therefore a collection of transmitter and receiver tunings. In this case the state generator outputs a sequence of mappings by assigning a wavelength for each path in the mapping. Each source node monitors the state for an assignment of a wavelength to its transmitter. When such a wavelength is assigned, the message is sent using the current mapping. Receivers must also monitor the state for wavelength assignments. In protocols which require the identification of the sender, they must also monitor the state to discover the transmitting node to which that wavelength was assigned. This is the general case, a number of simplifications are also possible. For example, either the transmitters (or receivers) could be assigned fixed wavelengths and only the receivers (or transmitters) could be tuned by the state generator.

A significant issue in sequence transformation for both of these examples may be latency of the transformation operation itself. For this reason we have also considered hierarchical algorithms. In these designs, a local algorithm makes rule based transformations in one step. Simultaneously, a global algorithm monitors overall trafffic and update the local rules accordingly.

Acknowlegments

This research has been supported in part by grants from the Air Force Office of Scientific Research, and the National Science Foundation, and by fellowship support from the Andrew Mellon Educational Trust.

1 References

[Dou91] P. Doud. High performance interprocessor communication through optical wavelength division multiple access channels. In *Proceedings of the 18th Symposium on Computer Architecture*, volume ACM 19(3), May 1991.

[Goo89] J. Goodman. Switching in an optical interconnect environment. In *OSA Proceedings on Photonic Switching*, volume OSA 3, March 1989.

Digital Optical Computing

OFB 10:30am-12:00pm
Grand Ballroom West

Frank A. Tooley, *Presider*
Heriot-Watt University, U.K.

Digital Optical Computing

K.H. Brenner
University of Erlangen-Nuremberg
Staudstr. 7
Erlangen, Germany D-8520

Summary not available at time of printing.

Spatial Soliton Dragging Gates and Light Bullets

Kelvin Wagner and Robert Mcleod
Optoelectronic Computing Systems Center
University of Colorado, Boulder CO, 80309-0425

Introduction

The asymmetric interaction between two orthogonally polarized, self-focusing spatial soliton filaments in a slab waveguide of nonlinear Kerr media can result in the spatial dragging of a strong pump beam by a weak signal wave. As a result, the strong pump is transmitted as a stable filament when the weak signal is not present, and in the presence of the weak signal wave it is dragged to the side by more than a beam width so that it can readily be blocked by an appropriate aperture. This dragging interaction is stronger than spatial soliton collisions,[1,2,3] and as a result absolute gain is achievable. High speed, phase insensitive, cascadable logic operations, such as NOR gates, can be realized with this technique by dragging a pump filament in successive stages by any of several inputs, just as in the Islam temporal soliton dragging gate.[4] But in this case the massive parallelism of free-space optical systems can be exploited, and enhanced nonlinearities can be utilized, dramatically decreasing the latency of the Islam gate.

Unfortunately, the self-focusing effect is not stable in two dimensions, and beam breakup is usually unavoidable during self-focusing of circularly symmetric beams, so slab waveguide confinement in one of the transverse dimensions is conventionally required in order to stabilize the filamentation. However, the self-focusing again becomes stable when the additional dimension of temporal pulse compression is added to the two transverse dimensions of self-focusing, resulting in stable "light bullets" in 3+1 dimensions.[5–7] These light bullets can act as the carriers of digital information and control signals, and can scatter off each other in spatial and temporal dragging interactions similar to the 1-dimensional gates. Multidimensional dragging allows the light bullet gates to perform logically complete operations such as NOR gates in a single stage. We are using these logical primitives to design a new class of 3-dimensional optical architectures that rely on the space-time flow of light bullets in analogy to 2-dimensional systolic arrays.

Filament Dragging

Two spatial solitons propagating at different angles can interact in a nonlinear medium resulting in a switching phenomena.[8] When two orthogonally polarized spatial soliton filaments collide in a nonlinear medium, each soliton shifts to the side without altering the original direction of propagation. Experimental demonstrations of spatial soliton collisions producing spatial shifts,[1] spatial soliton trapping,[2] and phase sensitive attractions or repulsions,[3] have recently been reported. These interactions have been suggested as mechanisms for a photonic switch, however no gain was demonstrated by any of these methods. This is because in these conventional symmetric collision, the effects before and after the collision are in the opposite directions and may nearly cancel, resulting in only small shifts of both spatial solitons.

In this paper, we investigate a different interaction geometry in which the two orthogonally polarized input beams are brought into coincidence at the boundary of the nonlinear slab waveguide. The intensity profiles overlap at the input to the nonlinear medium, but the two filaments are launched at different angles. By bringing the two beam profiles into coincidence preceding the nonlinear medium, the only force on the pump spatial soliton is to the left, and the signal filament is dragged towards the right, resulting in much larger shifts of both beams and the control of a large pump soliton by a smaller signal soliton which leads to gain. The two spatial solitons are now pulled towards each other

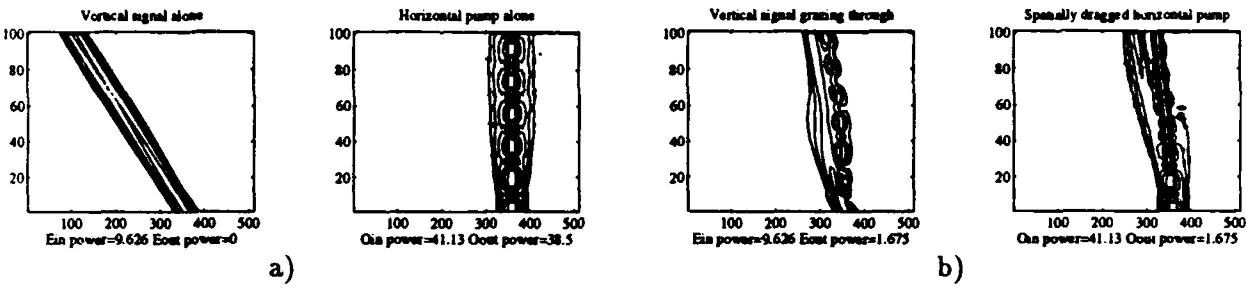


Figure 1: a) Independent propagation of a weak tilted e-polarized signal soliton and a strong o-polarized pump soliton, b) Simultaneous propagation of the two orthogonally polarized spatial solitons results in the dragging of the pump by the weak signal due to cross phase modulation.

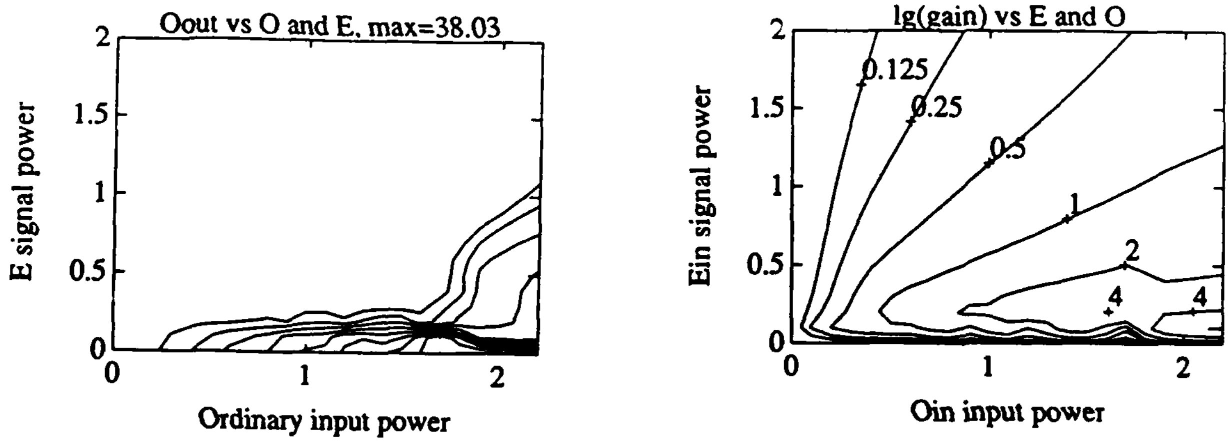
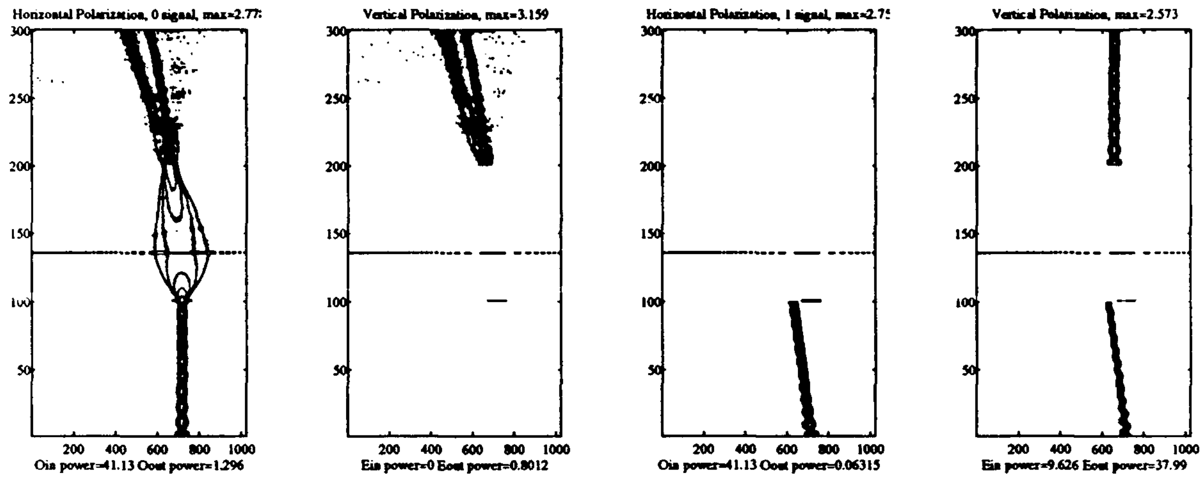


Figure 2: Pump output power passing through the output aperture, and resulting gain as a function of the signal and pump input powers.

so that they will orbit around each other's trajectory. Thus they continue to interact for the entire length of the nonlinear medium, resulting in large dragging of the pump by the signal, changing its angle of propagation in the nonlinear medium. The amount of dragging of one beam by the other is proportional to their relative power, when the powers are equal, it is to be expected by symmetry that the direction of propagation of the combined soliton pair will be the bisector between their initial directions of propagation. Gain is achieved by letting a small signal drag a large pump by more than a pump beam width so that it can be blocked by an opaque aperture. This is natural for this interaction geometry, since a small signal filament of one nth the power of the pump that is propagating at an angle that is m resolvable beam diameters away from the pump will drag the pump by $m/(n+1)$ beam widths. A simulation of this spatial dragging interaction is shown in Figure 1, where the weak orthogonally polarized signal beam can be seen to drag the strong pump by more than a beam width. This simulation showed an inverting gain of 4 and contrast ratio of 23:1. The limits on the gain are imposed by the minimal power that the signal beam can have and still form a stable soliton filament, and the maximum power that the pump can have and form a higher order spatial soliton that will not break up or induce other competing nonlinearities in its periodic high intensity foci (eg. Raman, Brillouin, 2-photon, etc). Absorption also limits the gain, while index saturation can increase the achievable gain by allowing the stable propagation of larger pump solitons. Simulations of this soliton



a) b)

Figure 3: Cascading of the output of one soliton dragging stage as the input to the next stage producing a logical restoration with gain. a) Zero input signal, b) One input signal.

dragging interaction for a wide variety of input signal and pump beam powers, as well as for saturating and nonsaturating nonlinearities, varying absorption, and different grazing angles have been performed, and the achievable gains, contrast ratios, throughputs and differential gains have been investigated and will be presented at the conference. An example of the pump output power passing through the output aperture and the achievable gain plotted as a function of the input power in the signal and pump is shown in Figure 2.

Cascading gates

The undragged pump from a spatial soliton dragging gate that passes through the output aperture can be imaged onto the input of a subsequent stage and used as its signal input. Beam propagation simulations of a cascade of two such soliton dragging gates with an intervening region of free space, and a holographic lens is shown in Figure 3. The second stage is pumped by an auxiliary beam with a polarization orthogonal to the pump in the first stage. Since the output pump that passes through the aperture is the logical inversion of the input signal, two stages of soliton dragging gates can be used as a logical signal restoration, or with feedback as a memory element. Similarly, the dragged output from the first stage can be used as the pump input to the second stage which can be dragged by the signal applied to the second stage, producing a phase insensitive logical NOR at the output aperture of the second stage.

Light Bullets

In media that have both a positive Kerr nonlinearity ($n_2 > 0$) and anomalous dispersion (negative group velocity dispersion), a pulse can collapse in both space and time forming a stable "light bullet" in 3+1 dimensions.[5–7] With different combinations of the signs of the nonlinearity and the dispersion, it is possible to construct dark bullets and other soliton solutions. These collapsed optical pulses have high peak powers but small total energies, making them attractive for computing applications. These solitons should be able to interact in the same way the one-dimensional spatial and temporal solitons do. In this case, however, there are three possible dimensions in which to drag (two transverse in space and one time dimension) resulting in a much richer framework for possible logic gates. A beam propagation simulation of a projection of a light bullet propagating in a linear and a non-linear media is shown in Figure 4.

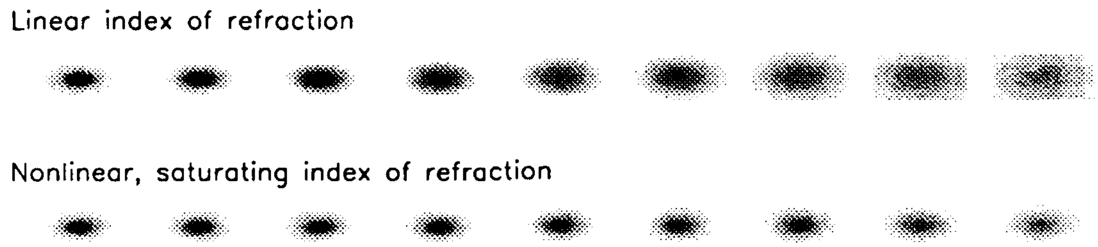


Figure 4: Comparison of linear versus light bullet propagation, calculated with beam propagation in the group velocity coordinate frame. The displayed fields are projections of the 3-D numerical solutions displayed to scale as a travelling wave movie. The initial condition is the theoretical 3-D spatial soliton eigenstate, and is only marginally stable because of the boundary absorptions.

The formation of a light-bullet approximately $5\mu m x 5\mu m x 10\mu m$ can be observed in the nonlinear media.

Architectures

The design of circuits and architectures for soliton logic is at an early state of development, although a decoder has been designed for the 1-D symmetric soliton collision gate.[8] Since the logical interaction between solitons can occur anywhere in the volume of nonlinear material where two beams cross, and not at specific gate sites, it is clear that circuit functionality is determined by the placing and timing of signals. With light bullets, dragging is possible in either of two space dimensions or time, and simultaneous dragging by two (or more) signals appears to be possible, allowing the compact single stage implementation of NOR gates. The primary difficulty in 3-D light bullet dragging computing may be getting the signals and pumps to the desired interaction sites without disruption by unwanted signals. We are investigating techniques to systolize complex interactions by pulsing the clock pumps and signals so that they pass through each other until arriving at the desired dragging logic site, and to program the functionality of the array of logic gates by the presence and absence of clock pumps in the space time lattice of possible light bullet locations.

Conclusion

A new type of ultrafast, low latency, all optical digital logic switch with gain, cascadability, input-output isolation, and phase insensitivity has been proposed and numerically demonstrated. This spatial soliton dragging gate and its extension to 3-D light bullet dragging opens up numerous new architectural possibilities for computing in 3-dimensions that may allow the realization of ultra high performance digital optical computers.

References

1. T. Shi and S. Chi, Nonlinear photonic switching by using the spatial soliton collision, *Opt. Lett.*, vol. 15(20), p. 1123, 1990.
2. M. Shalaby and A. Barthelmy, Experimental spatial soliton trapping and switching, *Opt. Lett.*, vol. 16(19), p. 1472, 1991.
3. J. S. Aitchison et al., Experimental observation of spatial soliton interactions, *Opt. Lett.*, vol. 16(1), p. 15, 1991.
4. M. Islam, All-optical cascadable NOR gate with gain, *Opt. Lett.*, vol. 15(8), p. 417, 1990.
5. Y. Silberberg, Collapse of optical pulses, *Opt. Lett.*, vol. 15(22), p 1282, 1990.
6. A.B. Blagoeva et al., Light bullets formation in bulk media, *IEEE JQE*, vol. 27(8), p 2060, 1991.
7. J. E. Rothenberg, Space-time focusing, *Opt. Lett.*, vol. 17(19), p 1340, 1992.
8. J.P. Robinson and D.R. Anderson, Soliton Logic, *Optical Computing and Processing*, vol. 2(1), p. 57, 1992.

A Case for All-Optical Digital Computing

Miles Murdocca
Department of Computer Science, Hill Center
Rutgers University, New Brunswick, NJ 08903
(908) 932-2654 (phone), (-0537 fax)
murdocca@cs.rutgers.edu

1. Introduction

Digital electronics is a relatively mature technology, and is approaching fundamental limits in terms of density and speed. Historically, there has been a hope that the use of optical logic gates and optical interconnects will improve these limits. Unfortunately, it is not certain that the use of optics at the gate level will improve density, and in fact, the packing density of optical logic gates may actually be less than the packing density of electronic logic gates due to diffraction and thermal limitations [1]. The greatest single limitation to packing density for electronics is the complexity of the interconnects, and this limitation may be relieved through the use of free space optical interconnects at the component level. Optical interconnects may thus improve system density for an opto-electronic computer, even if they do not offer an advantage at the gate level. Even at the gate level, however, optics can improve density in the sense that it is possible to achieve greater computing power with fewer logic devices through the use of reconfigurable interconnects.

A digital computer provides hardware for all operations that may be needed during the course of computation, even though only one operation is needed at a time. As a result, much of the hardware of a conventional computer is underutilized. If information is known about the hardware that a computation needs before it is used, then greater efficiency can be realized through a mechanism that reconfigures the computer during operation. The work reported here supports the opportunity to increase the performance of digital computers by reconfiguring gate-level optical interconnects to suit the changing needs of computations during execution.

2. Reconfigurable Logic

A potential improvement that reconfigurable optical interconnects offer that affects a broad range of processors is to tailor the size of the processor to match the changing needs of the computation being performed. Observations that most of the instructions that are executed in general purpose computers are simple (such as MOVEs) and involve few stack manipulations for subroutine linkage motivated the development of reduced instruction set computers (RISCs) such as the SPARC and RISC II. A related observation is that memory references tend to be localized, which led to the development of cache memories, in which the fraction of code that is active is kept in a small fast memory that is local to the processor. The result is a significant improvement in performance for a small increase in cost.

In addition to locality of reference which applies to memory references, there is also *functional locality* [1] which applies to instructions executing on a processor. The concept behind functional locality is that only a fraction of the hardware in a processor is used for a given interval of time, and that there is repetition within that interval. In support of this idea, we have taken measurements of instruction execution on running programs. What we have found is that even when a RISC processor with a small instruction set is used, that just a fraction of the available instructions are needed. For example, only 60 out of 180 available instructions might be used in a typical SPARC program. This is a previously known result. What we have observed that is significant here is that if we consider implementing just a few instructions at a time, say for example, 50% of the total number of different instructions needed in a program, then we can often achieve a hit ratio that is much greater than 50%, ranging between 85% and 97% for the samples we have taken. We take our measurements by generating a trace of instructions in the order that they execute, and then sliding a window of size N over the trace, which represents the N instructions that are simultaneously implemented in the processor at a given time, and then recording how many times an instruction is needed that is not currently implemented. The interconnection pattern for the instruction that generates the miss is then brought into the processor, at the expense of the interconnection pattern for some other instruction. As a result of these findings, the development of a *function cache* [2] has been proposed in which the most recently used hardware is kept in the form of interconnection patterns in a high speed section of a processor. A processor that supports functional locality must provide for gate-level reconfiguration, which motivates the case for gate-level optical computing as discussed in the next section.

3. An Optical Computing Model

A simplified form of the optical computing model considered here is shown in Figure 1, which originated at AT&T [3]. The model consists of arrays of optical logic gates interconnected in free space. The more general form of the model includes greater complexity in the logic arrays, such as electronic processing elements with optical I/O ports (smart pixels). For the simplified gate-level form, binary 1's and 0's are represented as intensities of light beams, and masks in the image planes block light at selected locations which customize the interconnects to perform specific logic functions. The system is fed back onto itself and an input channel and an output channel are provided. Feedback is imaged with a single row vertical shift so that data spirals through the system, allowing a different section of each mask to be used on each pass. The gate level interconnects are assumed to have a regular pattern such as a perfect shuffle or a crossover, which are used experimentally in a number of AT&T optical processor testbeds [3, 4] and in a processor under development in the Photonics Center at Rome Laboratory. The reason for using regular interconnects at the gate level is to allow the beams to share the same field of a single lens. If we choose to use irregular

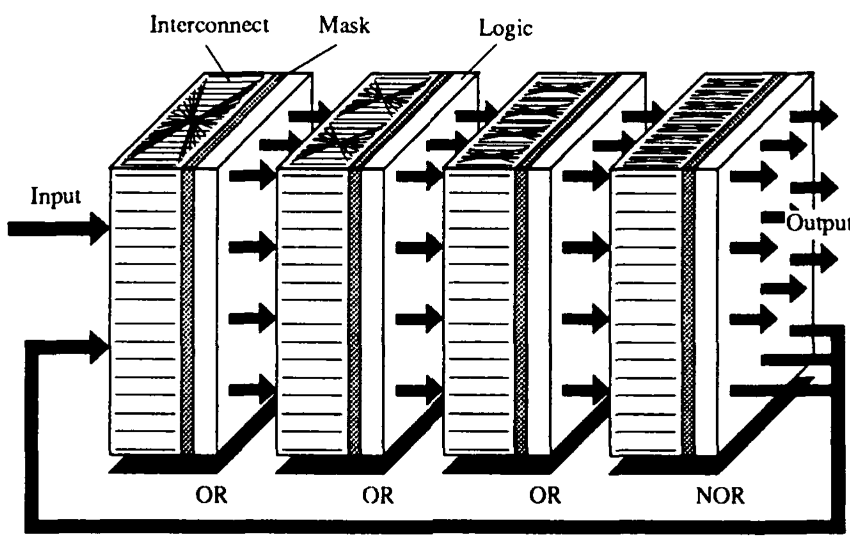


Figure 1: *The AT&T optical computing model.*

interconnects instead, then each channel requires a separate imaging system. This can be achieved through holography or through micro-optic fabrication techniques, but the small diameters of the resulting lenses limit propagation [1] due to diffractive coupling between neighboring channels. The two extreme cases that are characterized by trade-offs between regular and irregular interconnects do not necessarily exclude one another. In fact, a combination of the regular and irregular approaches, which are loosely representative of space-invariant and space-variant approaches, respectively, may support dense packing of devices while balancing the limitations imposed by practical realizations.

There is no need for the customizing masks to remain fixed, and in fact, they may be implemented in any of a number of ways, such as with ferroelectric liquid crystals, matrix addressable logic arrays, or through beam steering elements. In a very simple form, a reconfigurable logic/interconnection component (RELIC) consists of independently addressable optical logic gates and static optical interconnects. Reconfiguration at the gate-level is obtained by selectively enabling or disabling logic gates. Design methods based on this model have been previously developed as described in Ref. [5]. A RELIC may compete with existing electronic field programmable gate arrays (FPGAs) such as the Xilinx [6] line of reconfigurable components. The internal configuration of a Xilinx chip is shown in the left side of Figure 2. A number of programmable logic arrays (PLAs), shown as rectangles, are interconnected through an embedded arrangement of crossbar switches. The PLAs contain lookup tables (LUTs) for two seven-variable Boolean functions, and provide two bits of internal feedback to the LUTs. Each PLA generates two one-bit outputs. A small number of channels (five shown in the figure) pass through each crossbar in horizontal and vertical directions. The LUTs and crosspoints of the crossbars are configured by loading static flip/flops, one per decision element (or crosspoint). The Xilinx chips are popular in the area of rapid prototyping, in which a hardware implementation of a target processor is realized with reconfigurable components, but at a greater cost and with reduced performance than with a custom hardware design. Commonly, Xilinx chips are

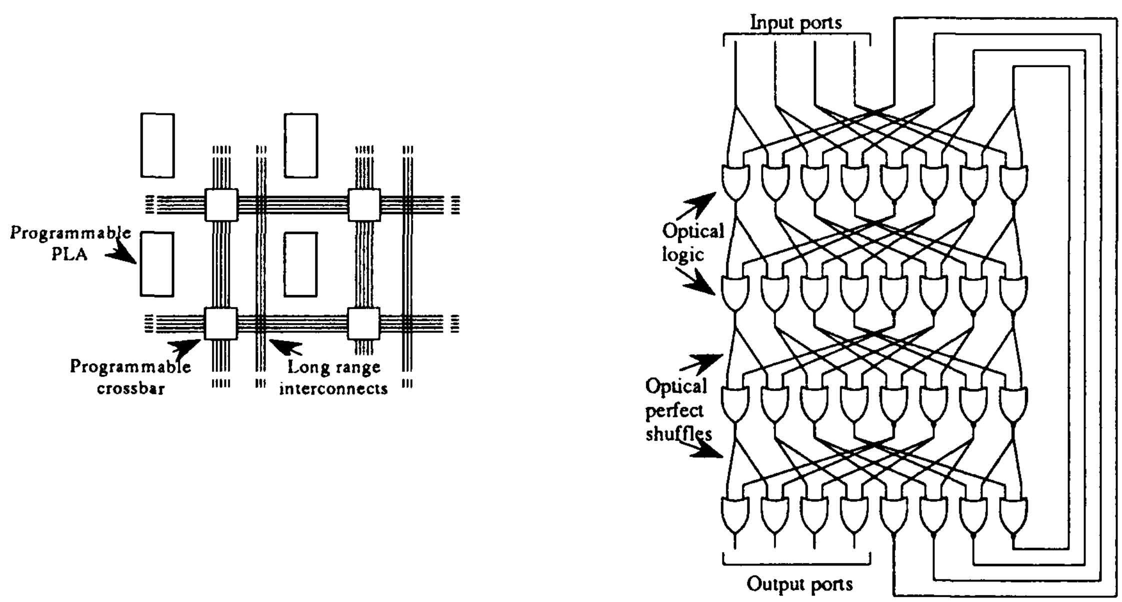


Figure 2: *Internal configuration of aXilinx chips(left); gate-level configuration of RELIC (right).*

used in end-products rather than just in transition hardware, particularly when production quantities are small (less than 1000 units) and when performance is not a significant goal.

With regard to reprogrammability, the Xilinx line is very flexible, but the user is forced to decompose large circuits into a number of interconnected one-bit circuits. This often unnatural decomposition sacrifices performance. For example, a ripple-carry adder maps well to the Xilinx approach, but a fast parallel adder does not. As an illustration of why this is the case, consider again the general layout of a Xilinx chip shown in the left side of Figure 2. The chip is clustered into one-bit logic units and narrow communication channels. Although it is possible to create a gate-level switching matrix that allows a user to modify interconnects at the gate or component level, it would be nearly impossible to maintain a clock speed of 50 MHz (a typical Xilinx internal clock speed) due to the enormous wiring complexity of such a chip. The RELIC approach allows gate-level and component-level interconnects to be allocated as needed, without causing a large increase in wiring complexity.

4. Conclusion

Anticipated benefits of reconfigurable interconnects include the development of self-modifying hardware that improves performance through use, and the ability of a computer to support complex operations in a relatively small volume. In order to exploit this capability, reconfiguration at the gate level is necessary, which may be achieved with all-optical logic.

The work reported here is jointly supported by the Air Force Office of Scientific Research and the Office of Naval Research under grant N00014-90-J-4018. The initial exploration that led to the function cache work was supported by the Strategic Defense Initiative Office under contract F49620-91-C-0055, which was administered through the Air Force Office of Scientific Research. The author expresses his thanks to Eugen Schenfeld at NEC for discussions on the Xilinx chips.

5. References

[1] D. Smith, M. Murdocca, and T. Stone, "Parallel Optical Interconnections," book chapter in *Optical Computing Hardware*, vol. 2, edited by S. Lee and J. Jahns, Academic Press, (1992), in press.

[2] M. Murdocca and V. Gupta, "Architectural Implications of Reconfigurable Optical Interconnects," to appear in *Journal of Parallel and Distributed Computing*, (Spring 1993).

[3] M. E. Prise, N. C. Craft, M. M. Downs, R. E. LaMarche, L. A. D'Asaro, L. M. F. Chirovsky, and M. J. Murdocca, "Optical Digital Processor Using Arrays of Symmetric Self-Electrooptic Effect Devices," *Appl. Opt.*, **30**, 2287, (1991).

[4] F. B. McCormick, F. A. P. Tooley, T. J. Cloonan, J. L. Brubaker, A. L. Lentine, R. L. Morrison, S. J. Hinterlong, M. J. Herron, S. J. Walker, and J. M. Sasian, "S-SEED Based Photonic Switching Network Demonstration," *OSA Proceedings on Photonic Switching*, H. S. Hinton and J. W. Goodman, eds., vol **8**, 48 (1991).

[5] M. Murdocca, *A Digital Design Methodology for Optical Computing*, The MIT Press, (1990).

[6] R. Freeman, "User-Programmable Gate Arrays," *IEEE Spectrum*, pp. 32-35, (Dec., 1988).

Optical Implementations of The Modified Signed-Digit Algorithm

Shaomin Zhou, Scott Campbell, and Pochi Yeh

Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106
Tel.(805)893-7015 Fax.(805)893-3262

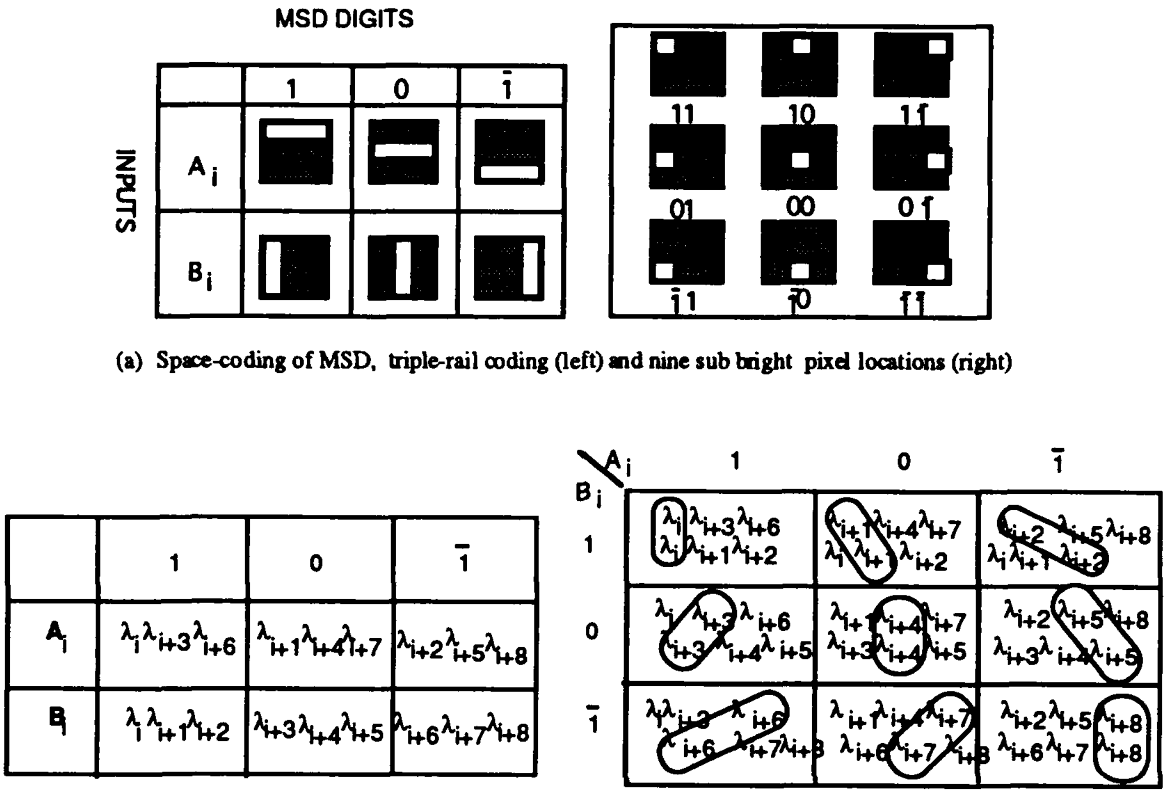
Hua-Kuang Liu

Center for Space Microelectronics Technology
Jet Propulsion Laboratory, California Institute of Technology
Pasadena, California 91109
Tel.(818)354-8935 Fax.(818)354-9509

Modified Signed-Digit (MSD) computing is an effective parallel processing method which is very suitable for optical implementation. The MSD arithmatic [1] was first introduced to optical computing by Drake et al [2]. Recently, several MSD related symbolic substitution approaches and higher-order MSD algorithms have been presented [3-9]. The symbolic substitution approaches offer an effective method to realize MSD computing and the higher-order MSD algorithms result in a reduction of the number of operation steps and an increase in processing speed. Although several optical systems based on linear or nonlinear optical elements have been proposed, many of them suffer drawbacks such as complicated optical set-ups or small spatial bandwidth products. In this paper, we present two novel optical implementations of the three-stage MSD algorithm. In the first approach, space-coding and multiple imaging via optical fan-out elements (OFE's)[10,11] are utilized. In the second approach, wavelength-coding [12] and four wave mixing [13] using photorefractive medium are utilized. Compared to other approaches, our first method has the advantages of larger throughput, higher operation speeds, and easier realization. As well, the advantages of our second method include simple structure and fiber parallel transportability of its 2-Dimensional signals.

A unique aspect of MSD is that it operates on a ternary numerical system composed of the numbers $\bar{1}$, 0, 1, thus making it carry-free [1]. For a three-step MSD algorithm, subtraction operations can be converted to addition operations by means of a complement code. Further, division operations can be converted to multiplication operations by means of parallel convergence division [3]. Finally, any addition operation can be implemented by using four types of transformation rules T, W, T', W' , and twice shift operations, and any multiplication operation can be implemented by using one type of transformation rule M and certain addition operations. Representations of the triple-rail MSD encoding in our approach are given in Figure 1. In (a), space-coding of A_i and B_i is utilized to represent the MSD values, and in (b) wavelength-coding of A_i and B_i is utilized to represent the MSD values (i denotes the digit location of the MSD's). In this fashion, one can formulate any one of $3\times3 = 9$ bright sub-pixel locations in the space-coding or 9 pairs of different wavelengths in the wavelength-coding by overlapping A_i and B_i. For n-digit MSD's operation, 9n sub-pixels in the space-coding or 9n different wavelengths in the wavelength-coding are required.

To implement the spatially coded optical MSD computations, we present the experimental set-up shown in Fig. 2. In this scheme, a plane wave is incident on overlapped input masks A and B which are the coded patterns of two operators according to the numerical representation system described in Fig. 1(a). The optical beam carrying the input signal then



(a) Space-coding of MSD, triple-rail coding (left) and nine sub bright pixel locations (right)

(b) Wavelength-coding of MSD, triple-rail coding (left) and nine pairs of wavelengths (right)

Fig.1 Coding approaches

propagates through a Dammann grating which serves as the OFE. The grating generates 3×3 angularly multiplexed copies of the input signal. Following this, the beam is split into three sub-beam paths using standard beam splitters (BS). In each of these three paths, the sub-beams are Fourier transformed and brought to pass through their respective Kernel filters (K_j, j=1-3). These Kernel filters act as spatial filters, selecting specific angular copies of the input signal. The spatially filtered beams are then allowed to propagate through decoding masks (D_j, j=1-3) which act as image-plane spatial filters the same as in the triple-in shadow-casting logical operation [14]. In this set-up, each sub-beam path processes a portion of the transformation (e.g. T), and the combination of these results provides the whole transformation (e.g. T). As mentioned, the various transformations are then utilized to realize various mathematical operations. The output of the system then contains a similarly coded information pattern found originally in either mask A_i or mask B_i, depending on the particular beam splitter alignments at the output port. Each of the transformations T, W, T', W', and M require three different Kernels (K_1, K_2, K_3) to complete the operations. These Kernels are placed in the OFE's Fourier plane and are given in Figure 3. The transmission of each of the Kernel elements determines which of the multiple images contribute to the final output. Figure 4 gives an example of the first stage of an addition operation in the three-step MSD algorithm. In Figure 4a are two inputs A and B. These would then be spatially coded according to Figure 1(a). The theoretical and experimental results of the first stage (the T and W operations) are then shown in Figure 4b and Figure 4c respectively. The second and third stages would then follow the algorithm as discussed in reference one. The coding information in the experimental results is chosen to be the same as that in mask A_i.

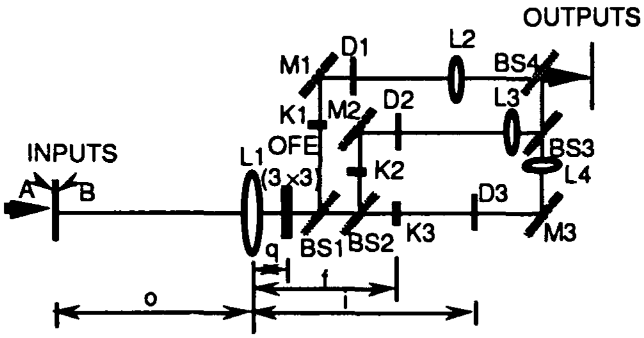


Fig.2 Optical set-up to implement space-coding MSD computing

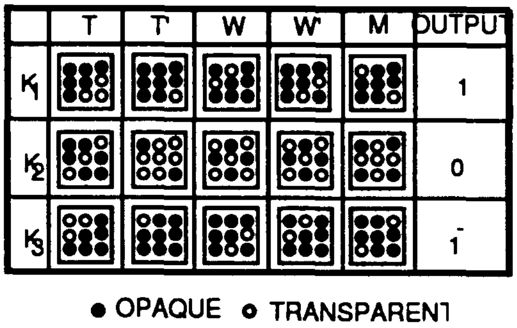


Fig.3 Kernel Configurations

An important system parameter to consider involves the achievable information throughput of the space-coding method. Information throughput will be affected by two different criteria. One is diffraction limitation, and the other is interference. These criteria are decided by the system parameters, such as the OFE dimensions, the optical wavelength, the object distance, the image distance, the distance between lens L_1 and the OFE, and the full beam angles between these three beams. For our set up, this value is about 10^5 bits per circle.

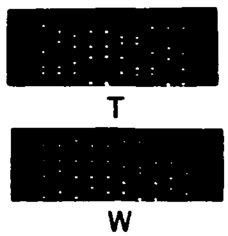
$$A=\begin{bmatrix}329\\-178\\-679\\689\end{bmatrix}_D=\begin{bmatrix}1&\bar{1}&0&1&1&\bar{1}&\bar{1}&0&0&1\\ \bar{1}&1&1&\bar{1}&0&1&0&\bar{1}&1&0\\ \bar{1}&\bar{1}&0&1&1&0&\bar{1}&0&1&\bar{1}\\ 1&1&0&\bar{1}&\bar{1}&1&0&1&\bar{1}&\bar{1}\end{bmatrix}_{MSD}\qquad B=\begin{bmatrix}550\\-319\\-402\\375\end{bmatrix}_D=\begin{bmatrix}1&1&\bar{1}&\bar{1}&\bar{1}&0&0&1&1&0\\ \bar{1}&1&\bar{1}&1&0&0&1&\bar{1}&\bar{1}&\bar{1}\\ \bar{1}&0&0&1&1&1&0&\bar{1}&1&0\\ 0&1&1&0&\bar{1}&1&1&0&0&\bar{1}\end{bmatrix}_{MSD}$$

(a) Two inpots

$$\begin{bmatrix}1&0&\bar{1}&0&0&\bar{1}&\bar{1}&1&1&1\\ \bar{1}&1&0&0&0&1&1&\bar{1}&0&\bar{1}\\ \bar{1}&\bar{1}&0&1&1&1&\bar{1}&\bar{1}&1&\bar{1}\\ 1&1&1&\bar{1}&\bar{1}&1&1&1&\bar{1}&\bar{1}\end{bmatrix}_{MSD}\qquad\begin{bmatrix}0&0&1&0&0&1&1&\bar{1}&\bar{1}&\bar{1}\\ 0&0&0&0&0&\bar{1}&\bar{1}&0&0&1\\ 0&1&0&0&0&\bar{1}&1&1&0&1\\ \bar{1}&0&\bar{1}&1&0&0&\bar{1}&\bar{1}&1&0\end{bmatrix}_{MSD}$$

T W

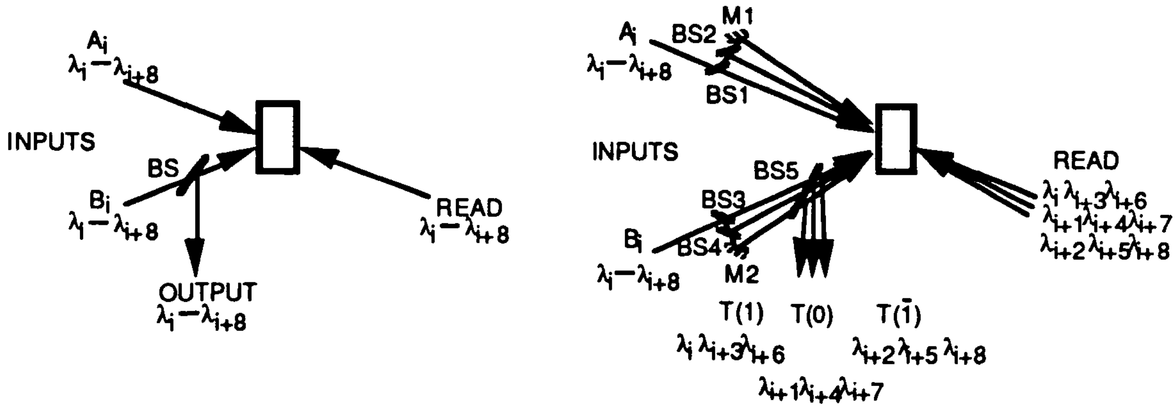
(b) Theoretical results



(c) Experimental results

Fig.4 A sample of space-coding MSD computing

The optical set-up to implement a wavelength-coded MSD algorithm is shown in Figure 5(a). Two multiwavelength inputs, A_i and B_i, coded according to the representation described in Figure 1(b) carry information via the presence or absence of light at the various wavelengths $\lambda_j \in (\lambda_i - \lambda_{i+8})$. These two beam sets interfere within a photorefractive medium and as a result, the wavelength present at one of $\lambda_i, \lambda_{i+1}, \ldots, \lambda_{i+8}$ will generate a grating corresponding to one of the input combinations 11, 10, ... , $\bar{1}\,\bar{1}$. The read beam set contains all of the wavelengths. Complete T (T', W, W', and M) operations in the three-step MSD algorithm may be realized by using the triple-in and triple-out set-up shown in Figure 5(b). Both inputs, A_i and B_i, are split into three paths via beam splitters,BS_1 - BS_4, and mirrors M_1 - M_2. The three pairs of beams of A_i and B_i write three gratings, G_1, G_0, and G_{-1} in the photorefractive medium. Consider, for example, an implementation of the T operation, as shown in Figure 5b. The three beam sets (λ_i, λ_{i+3}, λ_{i+6}), (λ_{i+1}, λ_{i+4}, λ_{i+7}), and (λ_{i+2}, λ_{i+5}, λ_{i+9}) are used to read the gratings G_1, G_0, and G_{-1}, respectively. As a result, the full T transformation signals will be obtained at the output port through beam splitter BS_5. This multiwavelength coded output may then be transported in parallel in an optical fiber, making its architecture compact and suitable for cascading.



(a) Single channel

(b) Three channels (full operation)

Fig.5 Wavelength-coding MSD computing

We have proposed two novel optical implementations of the MSD computing, the space-coding approach and the wavelength-coding approach. The experimental demonstration of the space-coding method by using OFE is presented.

References:

1. A. Avizienis, IRE Trans. Electron. & Comp. **EC-10**, 389 (1961).
2. B. L. Drake, R. P. Bocker, M. E. Lasher, R. H. Patterson, and W. J. Miceli, Opt. Eng. **25**, 38 (1986).
3. K. Hwang and A. Louri, Opt. Eng. **28**, 365 (1989).
4. Y. Li and G. Eichmann, Appl. Opt. **26**, 2328 (1987).
5. P. A. Ramanoorthy and S. Anthony, Opt. Eng. **26**, 821 (1987).
6. A. K. Cherri and M. A. Karim, Appl. Opt. **27**, 3824 (1988).
7. M. S. Alam, A. A. S. Awwal, and M. A. Karim, Appl. Opt. **31**, 2419 (1992).
8. M. M. Mirsalehi and T. K. Gaylord, Appl. Opt. **25**, 3078 (1986).
9. P. A. Ramanoorthy and S. Anthony, Opt. Eng. **26**, 821 (1987).
10. J. Tanida and Y. Ichioka, Opt. Lett. **16**, 599 (1991).
11. M. Wu, S. Zhou, J. Cai, and C. Cin, Opt. Commun. **86**, 454 (1991).
12. S. Maeda, T. Aoki, and T. Higuchi, Proc. IEEE 22st Int. Symp. on Multiple-Valued Logic, 282, June (1992).
13. Pochi Yeh, IEEE J. Q.E. **25**, 484, 1989.
14. S. Zhou, X. Yang, M. Wu, and C. Cin, Opt. Lett. **12**, 968 (1987).

Bit Serial Optical Computing

OFC 2:00pm-3:10pm
Grand Ballroom West

Matthew W. Derstine, *Presider*
Optivision

An Integrated Optics, Stored Program Computer

Harry F. Jordan, Alan R. Mickelson, Bart Van Zeghbroeck, and I. Januar
Optoelectronic Computing Systems Center, University of Colorado
Campus Box 525, Boulder, CO 80309-0525
(303) 492-7967

Introduction

A serial, stored program, optical computer has been designed [1] and implemented [2] using $LiNbO_3$ electro-optic directional coupler switches. Detector, amplifier and driver electronics at each electro-optic switch adapt it to optical control, so that all inter-switch signals are optical. The demonstrated machine uses discrete components, and glass fiber forms both interconnection and delay line memory. This paper will project the properties of an integrated version of this computer.

The integrated circuit uses $LiNbO_3$ for its optical section, including electro-optic switches, fixed ratio couplers, and memory registers, but excluding main memory, to be implemented off chip as an optical fiber loop. The photodetector, amplifier and electrode driver circuitry is built on a semiconductor chip which is flip-chip bonded to the $LiNbO_3$ using solder bump alignment [3]. Electrical connections from drivers to directional coupler electrodes are provided by the solder, while light is coupled from $LiNbO_3$ waveguides to photodetectors by surface gratings. Optical power is supplied by an external laser clock which generates pulses of the correct duty cycle at the bit rate. Figure 1 shows the $LiNbO_3$ waveguide structure for a 16 bit storage register with the schematic of the semiconductor electronics superimposed with dotted lines.

From known properties of the discrete component computer and from current research in integration technology, we will establish parameters such as clock rate, chip area, and power requirements for the integrated version. We will strike a balance between the conservative extreme of only using technology currently available off-the-shelf and the optimistic extreme of extrapolating current hero experiments ten years into the future.

System size

Implementation of main memory as a fiber loop implies an operating wavelength of 1300 nm as in the discrete component version, or perhaps 1500 nm. Either would satisfy the requirement for low loss and dispersion over the length of the fiber memory loop.

The serial optical computer required about 70 electro-optic directional couplers and about 80 fixed ratio couplers. Let the length of the coupling region be L and the waveguide convergence and

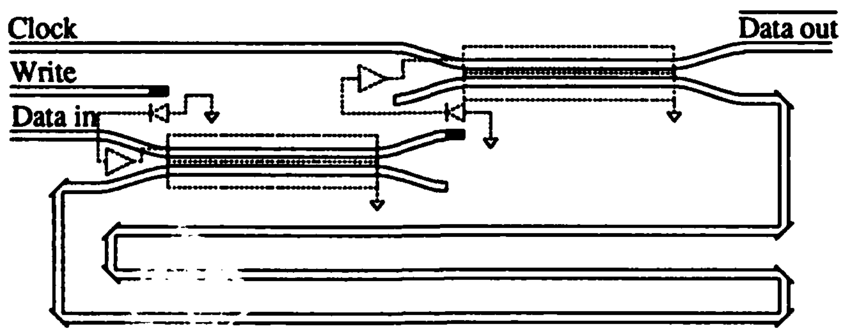


Figure 1: Memory register waveguides with overlaid electronics

divergence take distance M at each end. If the electrode structure requires a width E, the chip area for an electro-optic switch is $A_s = (L+2M)E$. Taking $L = 10$ mm, $M = 2$ mm and $E = 100\ \mu$m, $A_s = .014\ cm^2$, for about 70 switches per cm^2. Fixed couplers do not require electrodes, so taking their width as 30 μm, the area for a fixed coupler is $A_c = .0042\ cm^2$. 80 couplers would require .34 cm^2. These sizes are independent of the clock rate.

Since bits are stored in waveguide delays, a given capacity requires a given waveguide length and chip area. With a clock frequency of f bits/sec, an index of refraction of n, and speed of light, c, a bit occupies waveguide length $B = c/(fn)$. $LiNbO_3$ directional couplers have been demonstrated with a switching bandwidth of 40 GHz [4] and modulators up to 50 GHz [5]. If we take 20 Gbits/sec as an attainable bit rate and $n = 2.2$, then a bit occupies $B = 6.8$ mm in a $LiNbO_3$ waveguide. With a waveguide width of w and a separation of D between uncoupled guides, storing a bit for one clock period requires an area $A_b = B(D+w)$. $D = 20\ \mu$m and $w = 5\ \mu$m gives $A_b = .0017\ cm^2$ for about 588 bits/cm^2. The implemented design stored 142 bits in addition to main memory. If the waveguide can be suitably packed onto the chip, about one fourth cm^2 is required for storage. In addition to the above areas chip area must be allocated to interconnection and layout overhead. If this is less than .4 cm^2 the entire optical portion of the computer, exclusive of main memory, can fit within 2 cm^2 of $LiNbO_3$.

The main memory of 1024 sixteen bit words allowed by the prototype would require a fiber delay of 164 meters at 20 Gbits/sec. The integrated machine could easily add the 10 to 20 switches needed to break memory into four to eight loops for faster access. Multi-loop memories have been designed and simulated [6], but were not incorporated in the discrete component proof of principle machine.

With 2 cm^2 required by the optical circuit, the density required to implement 70 electronic circuits in a comparable area presents no fabrication problems.

Integration technology

Directional couplers on $LiNbO_3$ are a well understood technology, but folding long delay lines onto a small area requires discussion. The two competing technologies for folding are spiral with crossovers and corner mirrors. Bending loss measurements show that spiral waveguides in $LiNbO_3$ with a radius of curvature greater than 4 mm have negligible excess loss over straight waveguides [7]. A spiral of 1 mm diameter has been demonstrated in another materials system [8]. Crossovers required for a signal to exit from a spiral are not a problem. We have demonstrated that crosstalk is less than -30 dB in proton exchanged $LiNbO_3$ waveguides which cross at an angle greater than 6 degrees [9]. Corner mirrors have been demonstrated in other materials systems, e.g. [10], and it is predicted that a loss of .5 dB is attainable [11]. The difficulty of etching $LiNbO_3$ may make spiral folding preferable in the integrated computer. The low loss of all but very small angle crossovers gives great geometrical flexibility in laying out delay lines which cross signal paths and other delays.

The grating couplers, e.g. [12], needed to couple light to the detectors have negligible loss, and the dispersion angle is small enough that the ~30μm spacing of the solder bump bonding [3] allows the detector to collect virtually all of the light.

We choose a sub-micron-gate GaAs FET technology to implement the optoelectronic receiver, amplifier, pulse stretcher and switch driver. This technology has been shown to yield high speed optoelectronic receivers [13] as well as high speed amplifiers and multi-gigahertz logic. The detector favored is a Metal-Semiconductor-Metal detector which can be integrated with different GaAs FET technologies, while providing good responsivity, low dark current and high bandwidth [14] [15]. The GaAs FET technology has also been shown to be compatible with 1.3 μm wavelength by adding a GaInAs/GaAs superlattice absorbing region for the MSM detector [16] [17].

The projected receiver sensitivity is better than -20 dBm at a bandwidth of 20 GHz at the 1.3 μm wavelength. The MSM detector has a size of 30×30 μm and a responsivity of 0.2 mA/mW.

A transimpedance rather than an integrating receiver circuit is preferred for large bandwidth and high sensitivity. At least four amplifier stages are needed for sufficient current and voltage gain. The switch electrodes act as a transmission line with 50 Ω impedance. Allowing for amplifier power 40 mW of power are required for a 1 Volt switch. Since the solder bumps connecting the drivers to the switch electrodes are less than 30 μm in diameter, they do not influence the impedance calculation.

Power and losses

The system power is divided into two parts: optical power in the $LiNbO_3$ chip and fiber memory loops; and electrical power for the detector, amplifier and electrode driver electronics. Since drivers for about 70 switches in two cm^2 is fairly low density, the optical power budget is likely to be more important.

In the discrete component version both optical signal power and signal timing were restored by the technique of clock gating, as illustrated by the right hand switch in Fig. 1 used to regenerate register loop data. The discrete component machine required a signal restoring switch for any signal passing through the optical terminals of four others. In the integrated system, loss per switch is much lower since the signal does not pass in and out of fiber. This reduces both the optical power required from the optical clock and the number of switches needed. Not all restoring switches can be eliminated because any feedback loop must restore both power and timing to support recirculation. Losses in folding long waveguides must also be allowed for.

Dropping switches used strictly for power restoration, no optical path in the demonstrated machine passes through more than four 3dB couplers and six switched couplers. With .5 dB loss in an integrated switched coupler and its interconnecting waveguides, such a signal path would have an overall loss of 15 dB. As discussed above, we do not count the negligible loss in the grating couplers. The proof of principle system demonstrated reliable detection of -23 dBm signals photodetectors. If we estimate -17 dBm for the smallest signal at any detector in the integrated version, the maximum power needed at any clock input is -2 dBm. As discussed above, losses due to folding memory registers can be held to a few dB, and every such register includes signal regeneration, so these losses are not limiting.

The discrete component machine required 40mW of optical power. Scaling by the difference in the switch losses between the two versions gives about two milliwatts of optical power required from the laser clock. Another estimate takes the ten different clock copies required by the integrated machine at a worst case power of -2 dBm for a total power of about six milliwatts.

The optical power is all supplied by an off-chip optical clock generating pulses of about 50% duty cycle at the 20 GHz bit rate. A single mode locked semiconductor laser would easily satisfy this requirement, even with the worst case power estimate.

The electronics associated with each switch requires 40 mW for a total electronic contribution of 2.8 watts. The power can be easily dissipated with air cooling over the two square centimeter area which is determined by the optical chip.

Conclusions

This paper shows how an integrated optics, stored program, digital computer with a 20 GHz bit rate can be built using an optical system design demonstrated in a discrete component version and with optoelectronic integration technology currently available in the research laboratory. The flip chip bonded $LiNbO_3$ and $GaAs$ chips are only about two square centimeters in area and dissipate only a few watts of combined optical and electrical power.

REFERENCES

[1] V.P. Heuring, H.F. Jordan, and J.P. Pratt, "Bit-serial architecture for optical computing," *Applied Optics*, V. 31, No. 17, pp. 3213-3224 (10 June 1992).

[2] P.T. Main, R.J. Feuerstein, V.P. Heuring, H.F. Jordan, J.R. Feehrer, and C.E. Love, "Implementation of a General Purpose Stored Program Digital Optical Computer," submitted to Topical Meeting on Optical Computing, Optical Society of America, March 16-19, 1993.

[3] K. Katsura, T. Hayashi, F. Ohira, S. Hata, and K. Iwashita, "A novel flip-chip interconnection technique using solder bumps for high speed photoreceivers," *J. Lightwave Technology*, V. 8, No. 9 (Sept. 1990).

[4] S.K. Korotky, et al., "Optical intensity modulation to 40 GHz using a waveguide electrooptic switch," *Applied Physics Letters*, V. 50, pp. 1631-1633 (1987).

[5] D.W. Dolfi and T.R. Raganath, "Fifty GHz velocity matched, broad wavelength $LiNbO_3$ modulator with millimeter length active section," *Proc. Integrated Device Research Conf.*, PD2, New Orleans (April 1992).

[6] D.B. Sarrazin, "Optical delay line memory systems," *OCS Tech. Rept. 92-18*, Center for Optoelectronic Computing Systems, Univ. of Colo., Boulder, CO 80309-0525 (April 1992).

[7] I. Januar and A. Mickelson, "Low-loss PE:$LiNbO_3$ waveguide bends at λ = 1.55 μm," *Proc. Integrated Photonics Research Conf.*, New Orleans (April 1992).

[8] P. Verboom, Y.S. Oei, M. Smit, J. van Uffelen, H. van Brug, I. Moerman, G. Coudenijs, and P. Demeester, "A long InGaAsP/InP waveguide section with small dimensions," *IEEE Photonics Technol. Lett.*, V. 4, No. 10, pp. 1112-1114 (Oct. 1992).

[9] I. Januar and A.R. Mickelson, "Characteristics of S-shape waveguide structures by annealed proton exchange process in $LiNbO_3$," submitted to *J. Lightwave Technol.*, (1992).

[10] H. Appelman, J. Levy, M. Pion, D. Krebs, C. Harding, and M. Zediker, "Self-aligned chemically assisted ion-beam etched GaAs/AlGaAs turning mirrors for photonics applications," *J. Lightwave Technol.*, V. 8, No. 1, pp. 39-41 (Jan. 1990).

[11] D.K. Probst, L.G. Perrymore, B.C. Johnson, R.J. Blackwell, J.A. Priest, and C.L. Balestra, "Demonstration of an integrated, active 4×4 photonic crossbar," *IEEE Photonics Technol. Lett.*, V. 4, No. 10 (Oct. 1992).

[12] Q. Huang and H.J. Caulfield, "Waveguide holography and its applications," *Proc. SPIE*, V. 1461, Practical Holography V, SPIE Press, Bellingham, WA (1991).

[13] A. Ketterson, J.W. Seo, M. Tong, K. Nummila, D. Ballegeer, S.M. Kang, K.Y. Cheng, and I. Adesida, A 10 GHz bandwidth pseudomorphic GaAs/InGaAs/AlGaAs MODFET-based OEIC receiver," *Proc. Device Research Conf.*, Massachusetts, VIB-5 (1992).

[14] C.S. Harder, B.J. Van Zeghbroeck, M.P. Kesler, H.P. Meier, P. Vettiger, D.J. Webb, and P. Wolf, "High-speed GaAs/AlGaAs optoelectronic devices for computer applications," *IBM J. Research and Development*, V. 34, pp. 568,584 (1990).

[15] B.J. Van Zeghbroeck, W. Patrick, J.-M. Halbout, and P. Vettiger, "105 GHz bandwidth metal-semiconductor-metal photodiode," *IEEE Electron Dev. Lett., V. EDL-10, pp. 527-529 (1988).*

[16] A.N.M. Choudhury, C. Jagannath, A. Negri, B. Elman, and C.A. Armiento, "Thermally stable, superlattice-enhanced 1.3 μm InGaAs IMSM photodetectors on GaAs substrates," *IEEE Electron Dev. Lett.*, V. 12, p. 281 (1991).

[17] M. Zirngibl and M. Ilegems, "High-speed photodetectors on InGaAs/GaAs-on-GaAs superlattices," *J. Applied Physics*, V. 69, p. 8392 (1991).

Implementation of a General Purpose Stored Program Digital Optical Computer

P.T. Main, R.J. Feuerstein, V.P. Heuring, H.F. Jordan, J.R. Feehrer, and C.E. Love
Optoelectronic Computing Systems Center
Univ. of Colorado, Campus Box 525, Boulder, CO 80309
303-492-7967, 303-492-4293
email: toddm@fred.colorado.edu, heuring@fred.colorado.edu

Introduction

Optical computing has been a major field of research for a number of years with considerable effort expended on developing appropriate devices. A number of systems groups have adopted assorted techniques of dealing with the absence of good quality optical logic gates. Craig et al [1] have used the thermal index dependence of ZnSe nonlinear interference filters to implement slow all-optical logic gates and a small system with external control [2]. Two-dimensional arrays of SEEDs [3] have also been used to build a low speed processor with feedback [4]. The Programmable optoelectronic multiprocessor (POEM) uses VLSI circuits and silicon PLZT modulators to implement 2-D arrays of devices [5] . The digital optical computer of Opticomp [6], uses AND/OR logic at 100 MHz with external electrical data input, data output, and control, with no optical feedback. None of these systems implement a general purpose programmable optoelectronic computer with optical control. We report on the first such system in this paper.

A general purpose, bit-serial, stored program, digital optical computer [7] has been built and operated. Constructed as a proof-of-principle machine, it has all-optical signals between logic gates. The optical switches are housed in standard modules and are interconnected using optical fiber. This discrete component, fiber connected demonstration is extrapolated to an integrated optic version in a companion paper [8]. Unlike conventional electronic computers, synchronization is accomplished by time-of-flight synchronization [9], rather than by using latching elements. Information on how to design in this domain has been gathered from the experiment. Heuristics for time-of-flight synchronization were developed as timing problems arose and were overcome. Techniques for managing the complexity of the system were developed and refined as the project evolved. Characteristics required of the switching modules were observed, providing insight into future device requirements.

Goals

A key goal was to demonstrate a digital optical computer using all-optical signals between gates. Another was to use a latchless architecture which relies on time-of-flight synchronization. This timing technique is important to very high speed operation and is related to the wave pipelining technique [10]. A third goal was to design the computer with a stored program, general purpose architecture. A stored program, where both code and data reside in the same memory, is extremely important to the computer science community since it allows a computer to generate its own code using a compiler. A general purpose architecture implies that the instruction set is robust enough to allow execution of virtually any algorithm.

Optical Components

The components used in the computer are shown in figure 1. The optical splitter/combiner (a) is a passive device which combines the light entering the two inputs and splits it between the two outputs. It can be used to provide a passive-OR logic function and to fan out signals. It was experimentally confirmed that interference effects with the Fabry Perot laser diodes used was negligible. Single mode optical fiber (b) is used throughout the computer for interconnections and as a delay element.

The Lithium Niobate ($LiNbO_3$) directional coupler [11] has two polarization-sensitive optical inputs, two optical outputs, and an electronic control electrode (c). When a pulse is present at the control electrode, the switch enters the bar state, where light entering A exits through D (and B through E). Otherwise, the switch is in the cross state, where light entering A exits through E (and B through D). By connecting a photodetector and some pulse shaping electronics to the control electrode, the device can be used with all-optical inputs and outputs. A switch module contains six switches, their associated control terminal (C terminal) electronics, and polarization controllers. Eleven of these modules make up the computer.

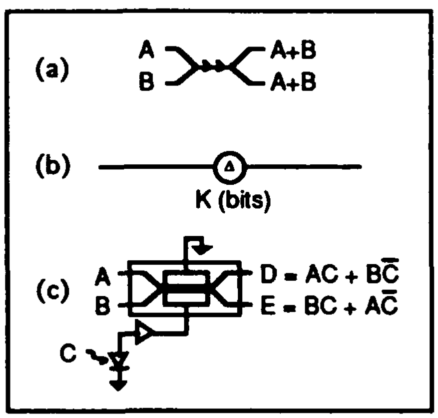


Figure 1. Optical Devices.

Architecture

The method used to provide memory (or *state*) is a recirculating delay line [12] (figure 2), where information is stored in space-time rather than statically in latches. Similar delay lines are used throughout the computer to implement registers and memory. Due to the cost of the optical switches, a bit serial design was chosen. The

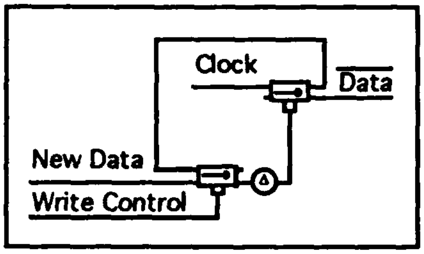


Figure 2. Delay Line Memory.

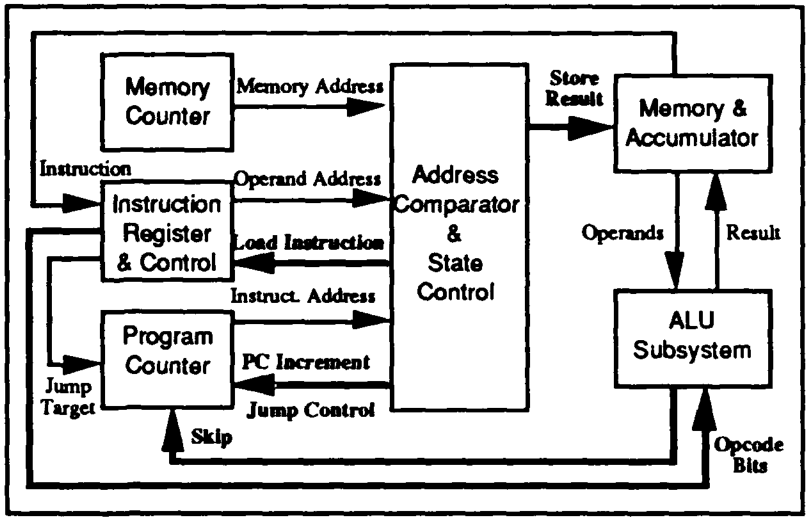


Figure 3. Simplified Block Diagram.

computer uses a 16 bit word size and a 50 MHz system bit clock, which is optically distributed. Since the goals of the project emphasize speed-of-light design rather than performance, a simple memory system was adopted, with sixty four data words [13,14].

As in other stored program computers, the operation of this machine relies heavily on the memory system. Each data word in memory recirculates serially in the delay line, passing the input and output ports once every memory cycle (a memory cycle is the time to complete one pass around the memory loop: 20.5 μs). A memory counter is used to mark the time when a particular memory word is available for reading and/or writing: the current value of the count is the address of the current memory word.

A block diagram of the ideal architecture (with no optical losses) is shown in figure 3, with major signal paths indicated (control signals are shown in bold face). It requires 54 switches, and was tested using an in-house simulator and CAD tool named XHatch [15]. The secondary architecture took device latency and loss into account, and included an additional 12 switches for amplitude and phase restoration. The XHatch model for this version took interconnection path lengths into account, and was used to calculate the lengths of the 200 optical fibers used in the computer.

The operations supported by the computer are shown in figure 4. Each instruction is executed in two steps. First, the machine compares the memory counter with the program counter (PC). When they match, the current memory word contains the desired instruction, which is loaded into the instruction register. Next, the memory count is compared with the address of the instruction's memory operand. When the operand is found, the result of the desired operation is computed and stored into either the memory or accumulator. If a jump is indicated, the PC is loaded with the address for the jump. If a conditional *skip* is indicated, the PC is incremented twice rather than once. Then the computer begins the search for the next instruction.

- Load
- Store
- Add
- Bitwise OR
- Bitwise AND
- Bitwise Complement
- Rotate Accumulator
- Jump
- Conditional Jump *(skip)*

Figure 4. Operations.

Implementation

The signals in the computer are return-to-zero encoded, with a 20% duty cycle. The maximum clock speed of the computer is determined by the bandwidth of the devices and the time to traverse the shortest feedback path. With the devices used, this is about 100 MHz [16]. A bit rate of 50 MHz was chosen for the computer since it is easier to maintain stable operation at this clock rate given the 175 MHz bandwidth of the switches used.

Some characteristics of the optical devices are loss, C terminal limits and crosstalk. The optical switches average about 5.5 dB loss each, while the splitters and fiber average about 0.3 dB excess loss per fiber connection. The minimum signal power that can be detected at the C terminal of a switch is about 5μW. A signal that will become attenuated beyond this limit must be restored, or *regenerated*, before it reaches 5μW by using it to switch a fresh copy of the clock in its place (figure 5).

In addition to the minimum power limit, a C terminal cannot detect a weak pulse of light that immediately follows a much larger pulse. If signals from different sources reach the same C terminal, it is necessary to attenuate one of the sources to match the two signal powers.

Crosstalk between switch outputs must also be controlled. When polarized properly, nearly all of the light entering an input will exit through the proper output according to the state of the switch (bar or cross). However, some of the light will exit through the other output as crosstalk, typically about -17dB.

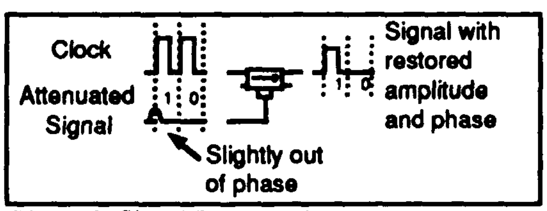


Figure 5. Signal Regeneration.

The time-of-flight nature of the architecture requires that precise timing and optical path length control be used to maintain synchronization. Since the method of timing restoration used is regeneration, proper clock timing is the key to maintaining system-wide synchronization. The most dependable way to keep the 18 copies of the clock synchronized is to use a single electronic clock signal to modulate a laser and fan out the optical signal using fiber. The system requires about 50 mW of optical power and uses four laser diode sources. This requires some electronic fanout of the clock, but most fanout was accomplished optically. Using one arbitrary copy of the clock as a reference, all other copies were adjusted until they matched the timing given by the XHatch model to within 0.1 ns.

The computer was built on a four level table, where the lower level houses power supplies, the next level houses clock lasers and clock fanout fibers, and the two upper levels house the switch modules. Individual subsystems were mapped from the XHatch circuits to the module layout with two goals in mind. First, all switches in a particular circuit were assigned to the same module whenever possible to reduce the length of interconnections. Next, subsystems which are heavily interconnected, such as the Memory, Accumulator and ALU, were placed as near one another as possible.

An electronic interface to the computer, the Monitor system [17], was used to help debug the optical circuits. It was implemented in two parts: a low-speed board provides an interface to a workstation, and four high-speed boards implemented in a high-speed ECL logic family interface with the C terminal electronics. The high-speed boards were mounted on the switch modules to provide the shortest possible path for the high-speed signals. As construction progressed, the Monitor became vital to the construction and debugging process.

Results

During the design and construction of the computer, synchronization problems were addressed at two levels: low level circuit timing and high level system timing. Low level synchronization was achieved by electronically stretching the pulse at a C terminal by about 4 ns. Arranging for this pulse to arrive 2 ns earlier than the other input(s) allowed timing variations of +/- 2 ns to be tolerated. In addition to circuit timing, the computer relies on a great deal of intercircuit connection and feedback. For instance, the PC is an input to the Address Comparator, which is an input to the State Control Logic, which returns to the PC. This large scale synchronization was accomplished by regeneration (figure 5). When a signal arriving at the C terminal of a regenerator switch was not synchronized with the clock at the other input, the timing of the C terminal signal was adjusted until it fully switched the clock. Because the clock signal had been carefully synchronized with the other copies of the clock, the signal exiting the regenerator was synchronized with the entire system.

As construction of the computer developed, much was learned about the characteristics of the devices. Although we implement pulse-stretching electronically, it could be done optically, at the cost of signal power. However, the high losses of the optical switches made this impractical. The use of integrated optics would greatly reduce switch losses, since most of the loss occurs at the fiber/$LiNbO_3$ junction. A significant problem that occurred during construction was electronic noise and crosstalk in the C terminal electronics. Integrating the electronics with the switches should greatly reduce this problem. A useful feature of the switches is the variety of logic functions they implement. This greatly reduces the number of gates needed to implement the computer.

Figure 6 shows an oscilloscope trace of an instruction fetch. At the top is the Memory-Found signal, which is generated when the desired memory word is available for reading. Because the machine state (INST/OP') indicates an instruction fetch, the Instruction-Found signal is generated. This signal is used to load the current memory word into the Instruction Register (IR). The bottom signal shows the IR, which reflects its new contents during the word period after the fetch. After this, the machine state changes to operand fetch in order to begin

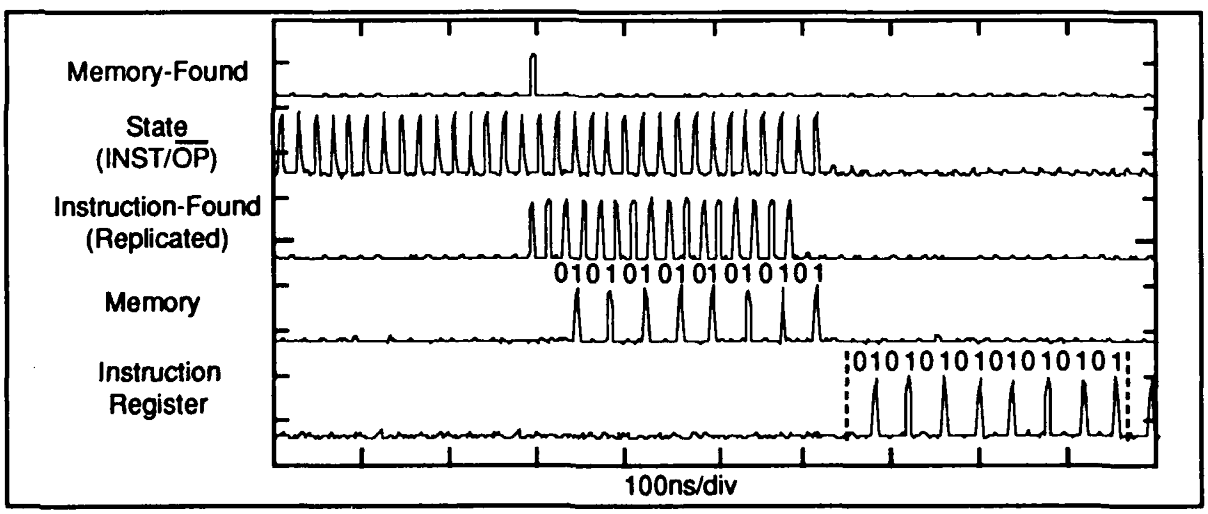


Figure 6. Oscilloscope trace of an instruction fetch.

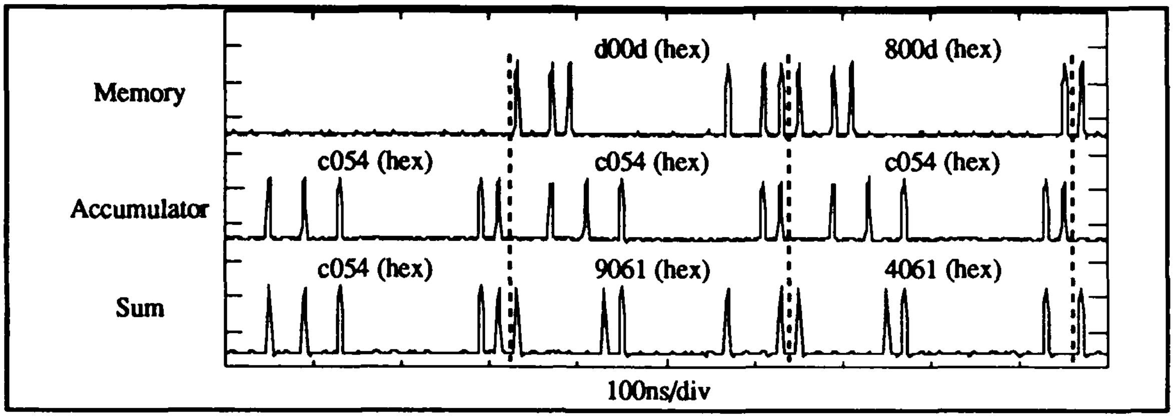


Figure 7. Addition of Memory and the Accumulator register.

execution of the new instruction.

Figure 7 is an oscilloscope trace of the output of the ALU full adder. The top two signals are the contents of the memory and accumulator, while the bottom trace is the sum of the two. Although this sum is computed every word period, it is only stored in the memory or accumulator when the opcode indicates an **add** instruction.

Conclusions

We have built and demonstrated a stored program digital computer in which all information is transmitted and stored as optical pulses. Photodetectors and electronic amplifiers are used with electro-optic directional couplers to perform logic. The experiment verified the effectiveness of time-of-flight, latchless, synchronization for high speed optical circuits. The machine was operated at a bit rate about half that implied by the bandwidth limitations of the directional coupler switches used.

References

[1] RGA Craig, GS Buller, FAP Tooley, SD Smith, AC Walker, BS Wherrett, "All optical programmable logic gate", *Applied Optics* 29(14) p. 2148 (1990).

[2] AC Walker et al, "Design and Construction of a programmable optical 16x16 array processor", Optical Computing Topical Meeting 1991 OSA Series Volume 6, paper TuA4-1 (1991).

[3] AL Lentine, DAB Miller, JE Henry, JE Cunningham, LMF Chirovsky, LA D'Asaro,"Optical logic using electrically connected quantum well pin diode modulators and detectors", *Appl.Optics* 29(14) p. 2153 (1990).

[4] M.E. Prise, N.C. Craft, M.M. Downs "Optical digital processor using arrays of symmetric self-electro-optic effect devices", *Appl.ied Optics* 30(17) p. 2287 (1991).

[5] F Kiamilev S Esener, VH Ozguz, SH Lee, "Programmable optoelectronic multiprocessor systems", in Digital Optical Computing, SPIE Critical Reviews Vol 35 RA Athale Editor, p. 197 (1990).

[6] P.S. Guilfoyle, Digital Optical Computer Fundamentals, Implementation, and Ultimate Limits", in Digital Optical Computing, SPIE Critical Reviews Vol 35 RA Athale Editor, p.288 (1990).

[7] V.P. Heuring, H.F. Jordan, J.P. Pratt, "Bit Serial Architecture for Optical Computing", *Applied Optics* 31, pp 3213-3224 (1992).

[8] H.F. Jordan, A.R. Mickelson, B. Van Zeghbroeck, I. Januar, "An Integrated Optics, Stored Program Computer", submitted to Topical Meeting on Optical Computing, OSA, March 16-19, 1993.

[9] J.P Pratt, V.P. Heuring, "Delay Synchronization in time-of-flight optical systems", *Applied Opt*ics 31, pp. 2430-2437 (1992).

[10] D. Wong, G. De Micheli, M. Flynn, *Proc. IEEE ICCAD '89*, pp. 270-273 (1989).

[11] F.T. Stone, J.E. Watson, D.T. Moser, W.J. Minford, SPIE Vol. 1177, p. 322, (1989).

[12] D.B. Sarrazin, H.F. Jordan, *Applied Optics* 29, 627-637 (1990).

[13] T.J. Soukup, R.J. Feuerstein, V.P. Heuring, "Implementation of a fiber optic delay line memory", *Appl.ied Optics* 31, pp 3233-3240 (1992).

[14] D.B. Sarrazin, "Delay Line Memory Systems", OCSC Tech. Rept. 92-18 (1992).

[15] M. Salerno, "X-Hatch Users Manual", OCSC Tech. Rept. 91-25 (1991), University of Colorado, Boulder.

[16] R.J. Feuerstein, T. J. Soukup, V.P. Heuring, "100-MHz optical counter that uses directional coupler switches", *Optics Letters* 16, 1599-1601 (1991).

[17] J.R. Feehrer, "BMON-II: A New Design for the BSOC Monitor", OCSC Tech. Rept. 92-02 (1992), University of Colorado, Boulder.

OCSC Technical Reports are available upon request.

Time of Flight Packet Synchronizers

Carl E. Love and Harry F. Jordan
Optoelectronic Computing Systems Center, University of Colorado,
Campus Box 525, Boulder, CO. 80309-0525.
Email love@fred.colorado.edu, harry@fred.colorado.edu

Introduction

The data rates of modern communication networks are reaching speeds at which traditional flip-flops are not effective. The problem is that the bit time is on the same order of magnitude as the feedback time of the flip-flop. In this region of operation the flip-flop no longer exhibits the expected static behavior. Furthermore the data rates are assumed to be as high as possible so that we can not use a local clock, running at a multiple of the bit rate, to measure the arrival skew of packets. The application we are specifically targeting at is a routing node within an optical communication network. The routing nodes have two inputs and two outputs. The routing node must examine the addresses contained in arriving packets to decide how to route them. The packets must be aligned so that the corresponding address bits can be examined in real time. The purpose of this paper is to look at two approaches to building circuits which can synchronize two incoming bit serial optical packets without the use of flip-flops.

The design technique which will be used in the two packet synchronization methods presented in this paper is referred to as time of flight design [4]. Time of flight design is based on the assumption that the propagation time of signals between and through logic gates is constant and known. The interconnection paths are adjusted so that the signals arrive at the logic gates at the same time. The precise controllability of optical delays makes this possible for optical networks.

The two solutions to be presented assume that there exists a control signal which is a logic 1 for the entire length of the packet. The control signal is examined to determine if a packet is at the node. The control signal simplifies the design by making the packet detection and alignment independent of the contents of the packet.

Linear Delay Line

The Counterpropagating delay line solution consists of a linear sequence of interconnected switching nodes [1]. The packets traverse the nodes in opposite directions. Figure 1 shows the basic architecture of the line with two packets which have just entered a line of equally spaced switching nodes. Since the packets are counterpropagating they will eventually collide at some node i assuming the skew between the leading edges of the packets is less than the length of the line. Figure 2 shows the two packets being routed out of the line and into the output network after they collided at node 4. The node passes an incoming packet to the next node if only one packet is seen approaching the node. If the node i sees two packets approaching which will collide at node i then it will change to the exit state and route the packets out of the line. Node i decides if it should be in the pass state or the exit state by examining the status signals from node i and its neighbors. Each node has two status signals, P and Q. The status signals are available to the adjacent nodes with minimal delay while the data packets suffer a delay between nodes. The value of Q will be a 1 if there is a

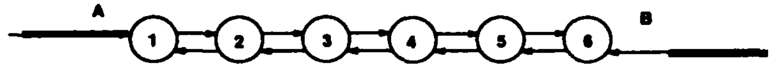


Figure 1: Linear Delay Line

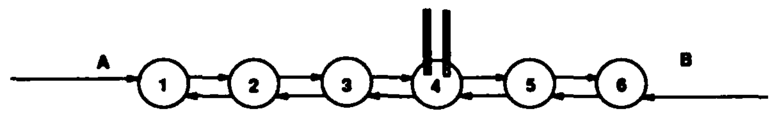


Figure 2: Linear Delay Line

right-to-left traveling packet at the node, it will be zero otherwise. Similarly, the value of P will be a 1 if there is a left-to-right traveling packet at the node, it will be zero otherwise. Each node can see the value of P for the node to its left and the value of Q on its right. Node i will change to the exit condition if it sees $P_{i-1} = 1$, $P_i = 0$, $Q_i = 0$, $Q_{i+1} = 1$. It will stay in the exit state until P_{i-1}, P_i, Q_i, Q_{i+1} are all zero. Note several assumptions are made about the packets and the P and Q signals: a packet is at least as long as the internode spacing; the spacing of the packets on each line is at least equal to the node spacing; the control circuitry assumes fundamental mode operation; and the signal propagation time can be very accurately controlled. Figure 3 shows the detailed line architecture. The figure also shows the output network. The outputs from each of the routing nodes for the left-to-right traveling packet are combined into a single output. Similarly the right-to-left traveling packet outputs are combined into one output. The alignment of the packets is maintained by the output network.

Fundamental mode design techniques were used to derive the logic equations for the node state. The line was then designed with Lithium Niobate 2x2 crossbar switches [5] and fiber optic interconnections. The set of all possible input status signals for a given node were examined to ensure that only one node at a time would see the exit condition. From the set of input signals a reduced state diagram was found. The state diagram consists of three states: the node is in the exit state; the node is in the pass state; and the node is idle state waiting for input. The design of the delay line was simulated on an optical time of flight circuit simulator, XHatch, developed at the University of Colorado [3].

This approach allows two packets to be aligned to a tolerance which is at least as small as the switching node spacing, assuming the input skew is less than the length of the line. The

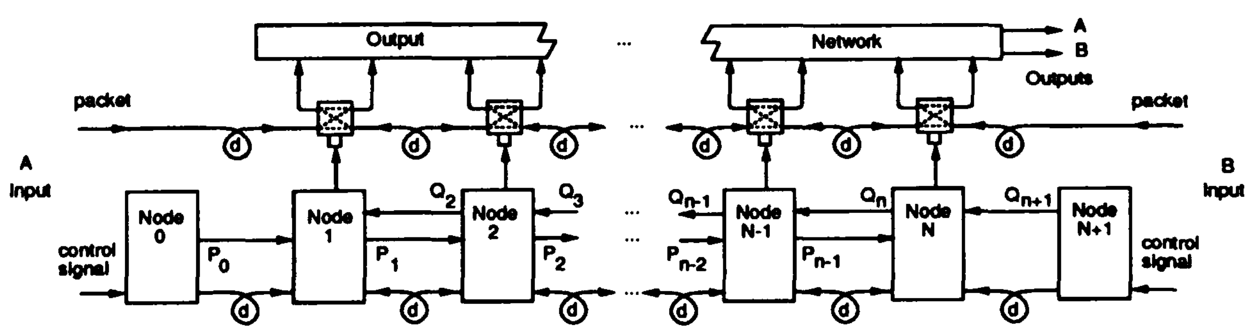


Figure 3: Detailed Linear Delay Line

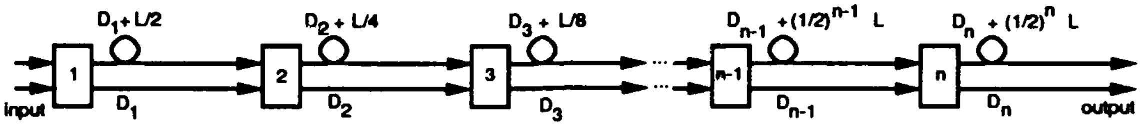


Figure 4: Logarithmic Delay Line Structure

spacing of the packets on each of the input lines is equal to the node spacing. Finally, the number of nodes required is proportional to the alignment tolerance and the packet length.

Logarithmic Delay Line

The logarithmic delay line consists of a sequence of switching nodes in which two input packets travel in the same direction [2]. The nodes are connected by two paths whose lengths differ by a fraction of the packet length. Figure 4 shows the logarithmic delay line architecture. The difference in the path lengths between node i and node $i+1$ is given by $(1/2)^i L$, where L is the length of a packet. The longer of the two lines is referred to as the delayed line, the other line is referred to as the nondelayed line. The switching node routes the first packet to appear on either of its inputs to the delayed output line. The maximum skew of the input packets is halved at the input to each succeeding node. The following assumptions about the packets have been made: the packet length L is fixed; the packets on the same line are separated by at least L; the design of the switching node circuitry is based on fundamental mode operation; there is a control signal associated with each packet; and the signal propagation can be accurately controlled. The proof that the maximum input skew is halved at the input to successive nodes is given in the following theorem.

Theorem:

The lengths of the input packets are assumed to be equal and fixed. Given that the skew at node i is bounded between 0 and $(1/2)^{i-1}L$, where L is the length of a packet then the skew at node $i+1$ will be between 0 and $(1/2)^i L$.

Proof:

Let t_i^l be the time the leading packet entered the delayed line at node i and let t_i^t be the time the trailing packet entered the nondelayed line. Define t_{i+1}^0 be the arrival time of the packet at node $i+1$ on the delayed line and let t_{i+1}^1 be the arrival time on the nondelayed line. The skew at node $i+1$ is given by:

$$| t_{i+1}^1 - t_{i+1}^0 | = | t_i^t + D_i - t_i^l - D_i - (1/2)^i L | \quad (1)$$
$$| t_{i+1}^1 - t_{i+1}^0 | = | t_i^t - t_i^l - (1/2)^i L | \quad (2)$$

The value of $t_i^t - t_i^l$ is bounded between 0 and $(1/2)^{i-1}$, thus from equation 2 the bound at node $i+1$ is:

$$0 \leq | t_{i+1}^1 - t_{i+1}^0 | \leq (1/2)^i L \quad (3)$$

The design of the switching node logic is based on two signals P and Q. The value of P is a logic 1 if there is a packet on the delayed input to the node, it is a logic 0 otherwise. Similarly the value of Q is a logic 1 if there is a packet on the nondelayed input, zero

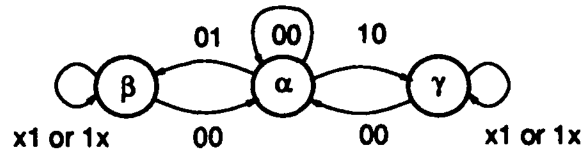


Figure 5: Reduced State Diagram

otherwise. If the values of P and Q are 0 and 1, respectively, then the node routes the packet on the delayed input to the delayed output. If P and Q are 1 and 0, respectively, then the nondelayed input is routed to the delayed output and the delayed input is routed to the nondelayed output. If the values of P and Q are both ones, then the node has already made a routing decision which will be maintained until both inputs return to zero. The reduced state diagram for the node control logic consisted of three states: the idle state (α) where P and Q are zero; the exchange state (γ) where the delayed input is connected to the nondelayed output; and the nonexchange state (β) where the delayed input is connected to the delayed output. Figure 5 shows the reduced state diagram.

The control circuits were derived using fundamental mode analysis. The design was simulated using XHatch to verify the correctness of the architecture. The number of nodes required is logarithmic with respect to the ratio of the packet length to the desired alignment tolerance. The spacing of the packets on the input lines must be at least equal to the packet length. The alignment tolerance is only a function of the minimum difference in the internode paths.

Conclusions

Two techniques for synchronizing packets of information have been shown. These techniques do not use traditional flip flop elements. The first approach required a linear number of nodes with respect to the alignment tolerance and the packet length. The second approach requires a logarithmic number of nodes with respect to the alignment tolerance and the packet length. The spacing of the nodes in the second approach is not important, only the difference in the lengths is important. The two approaches have been simulated to verify their operation. The simulations were done assuming an optical implementation because the signal propagation can be accurately controlled in optics.

References

[1] Carl Love, "A Time of Flight Packet Synchronization Technique", Report 92-19, Optoelectronic Computing Systems, University of Colorado, Boulder, CO. 80309.

[2] Carl Love, "A Logarithmic Time of Flight Packet Synchronization Technique", To be published, Optoelectronic Computing Systems, University of Colorado, Boulder, CO. 80309.

[3] Marco Salerno, "XHatch, User's Manual", OCS Tech Report 91-25, Optoelectronic Computing Systems, University of Colorado, Boulder, CO. 80309.

[4] Jonathan Pratt and Vincent Heuring, " Delay Synchronization in Time-of-flight Optical Systems", Applied Optics, Vol. 31, No. 14, May 10, 1992, pp. 2430 - 2437.

[5] Robert Feuerstein, "Lithium Niobate Switch Evaluation", OCS Tech Report 90-16, Optoelectronic Computing systems, University of Colorado, Boulder, CO 80309.